## REMARKS

Claims 27, 29-42, 44-47, and 49-55 are pending in the present application. Reconsideration is respectfully requested in view of the following comments and the amendments above.

The rejection of Claims 27-29, 35, 39, 43, 44, and 48 under 35 U.S.C. §112, first paragraph (written description; new matter), is obviated by the present amendment.

Applicants make no statement in regard to the propriety of this ground of rejection and in no way acquiesce to the same. However, in order to expedite examination of this application, Applicants have deleted the phrase "all or part of" from Claim 27. Therefore, this ground of rejection is now believed to be moot.

Withdrawal of this ground of rejection is requested.

The rejection of Claims 27-29, 35, 39, 43, 44, and 48 under 35 U.S.C. §112, first paragraph (written description and enablement), is obviated by the present amendment.

In regard to the Examiner's criticisms of the claims in relation to SEQ ID NO: 11, Applicants note that this issue is now moot in view of the deletion of this sequence by amendment herein.

Turning to the claims as currently presented, Applicants wish to offer the following remarks. In particular, to address the Examiner's assertion that "extensive random unpredictable experimentation would be required to determine useable peptide molecules for a given purpose," Applicants offer the following:

MPEP §2164.04 states:

> A specification disclosure which contains a teaching of the manner and process of making and using an invention in terms which correspond in scope to those used in describing and defining the

> subject matter sought to be patented must be taken as being in compliance with the enablement requirement of 35 U.S.C. 112, first paragraph, unless there is a reason to doubt the objective truth of the statements contained therein which must be relied on for enabling support.

Applicants submit that the present specification provides ample description and support for the skilled artisan to produce polypeptide molecules containing at least 50 consecutive amino acids of SEQ ID NO:3, since the amino acid sequence of SEQ ID NO: 3 is provided in the present specification. Further, the skilled artisan may, without undue experimentation, easily synthesize at least 50 consecutive amino acids according to the methods described at least on page 11 of the specification; i.e., by the methods of Houber-Weyl or R.D. Merrifield. These methods have been known in the art for at least 31 years!

Following synthesis, the particular peptides can be tested for their ability to confer immunity against malaria as described throughout the Examples in the present specification. More particularly, the skilled artisan would initially, after synthesis, try to test the peptides in mice as illustrated in Example 6 of the specification. After initial screening, the peptides can be further tested in chimpanzees as exemplified in Examples 2 (2.1 to 2.5), 4 (4.1 to 4.4) and 7. Thus, there are numerous working examples in the present specification that would guide the person skilled in the art to test for immunogenicity.

Furthermore, the specification provides additional guidance with respect to the particular peptides that should be chosen. For example, the specification describes that B and T responses are important and have been identified in the non repeat proteins of the 5' and 3' regions of LSA-3 (page 19). At least Example 3 of the specification describes ways to identify CTL epitopes and that observed protection that depended on cellular responses is important for a vaccine.

Moreover, on page 35 the specification describes that inhibitory activity is displayed in the development of *Plasmodium falciparum* with interferons. Thus, the person skilled in the art can also test to determine whether the peptides have this type of activity.

Thus, Applicants submit that the specification provides enough guidance to permit the skilled artisan to obtain and test the polypeptide molecules containing at least 50 consecutive amino acids of SEQ ID NO: 3 without undue experimentation and, as such, satisfy the requirements of MPEP §2164.04.

Applicants caution that the quantity and complexity of testing should not be a criteria to render a rejection under enablement. As set forth in *Massachusetts Institute of Technology v. AB Fortia*, 774 F. 2d 1104, 227 USPQ 428 (Fed. Cir. 1985):

> [T]he fact that experimentation may be complex...does not necessarily make it undue, if the art typically engages in such experimentation.

In fact, MPEP §2164.06 states:

> ... quantity of experimentation needed to be performed by one skilled in the art is only one factor involved in determining whether "undue experimentation" is required to make and use the invention. "[A]n extended period of experimentation may not be undue if the skilled artisan is given sufficient direction or guidance." In re Colianni, 561 F.2d 220, 224, 195 USPQ 150, 153 (CCPA 1977). " 'The test is not merely quantitative, since a considerable amount of experimentation is permissible, if it is merely routine, or if the specification in question provides a reasonable amount of guidance with respect to the direction in which the experimentation should proceed.

Applicants submit that, with the present specification in hand, determination of polynucleotide sequences that fall within the scope of the present invention would require nothing more than routine experimentation, even though the protocols may appear to be complex. As such, Applicants submit that the claims of the present application are fully enabled within the context of 35 U.S.C. §112, first paragraph.

Based on the foregoing, Applicants submit that the present claims are fully described and enabled by the specification, especially when viewed in conjunction with the common

knowledge available in the art. As such, withdrawal of these grounds of rejection is requested.

The rejection of Claims 27-29, 35, 39, 43, 44 and 48 under 35 U.S.C. §112, second paragraph, is obviated by amendment.

Applicants have amended the claims to be free of the Examiner's criticisms. More specifically, Claim 27 has been amended to recite at least 50 amino acids of SEQ ID NO: 3. It should be clear from the language in this claim, that any polypeptide molecule having one or more polypeptides of SEQ ID NOS: 10, 12, 13, 14 and 15 are *in fact* excluded from the claim language. Since "all or part" has been deleted from this claim, it is clear that peptides shorter than at least 50 consecutive amino acids are not encompassed by the claim. Moreover, Claim 35 has been amended to ensure proper antecedent basis.

In view of the amendments herein, withdrawal of this rejection is respectfully requested.

The rejection of Claims 27-29 under 35 U.S.C. § 102 (a) over Barnes et al is traversed.

The Examiner asserts that the argument filed on September 3, 2004, has been considered but is not persuasive. The Examiner further asserts, "Applicant cannot rely upon the foreign priority claim to overcome this rejection because the requirements of 35 U.S.C. §119 have not been met in... the instant application. The certified copy of the priority document is absent."

Applicants submit that this assertion by the Examiner is incorrect in view of the record of the present application. On December 22, 2000, Applicants filed a Request for Priority (copy **submitted herewith**), in which Applicants claimed priority to FR9507007.

Further, in the Request for Priority, Applicants noted that a certified copy of the same was submitted to the International Bureau in PCT application PCT/FR96/00894. It was further noted that receipt of the certified copy by the International Bureau in a timely manner under PCT Rule 17.1(a) was acknowledged as evidenced by PCT/IB/304 (copy **submitted herewith**).

Applicants again note that Barnes et al was published in August 1995, whereas this application claims priority to French application 95/07007 filed on June 13, 1995. Since the certified copy of FR9507007 was timely submitted to the International Bureau in PCT application PCT/FR96/00894, a certified English translation of this French application was filed in the parent application (US 08/973,462), and a certified English translation of this French application was again filed with the response on September 3, 2004 (copy of the date-stamped filing receipt is **submitted herewith**), the Examiner's rejection is not proper. Nonetheless, in order to ensure that no further errors occur in this regard, Applicants **enclose herewith** *another* copy of the certified English translation of the French priority document (FR9507007). Accordingly, Applicants submit that Barnes et al should not be prior art and this ground of rejection should be withdrawn.

Acknowledgment of withdrawal of this rejection is respectfully requested.

The rejection of Claims 27-29, 35, 39 and 43 under 35 U.S.C. § 102 (f), is respectfully traversed.

In rendering this rejection, the Examiner deems that since U.S. Patent Nos. 6,319,502 B1 and 6,270,771 have different inventors than the presently claimed invention and describe similar subject matter, the presently claimed inventors did not contribute to the above-captioned invention. Applicants disagree with the Examiner's conclusions for the reasons already of record.

It should be noted that Pierre Druihle is a common inventor for both the U.S. issued patents and the present invention. Furthermore, U.S. Patent Nos. 6,270,771 B1 and 6,319,502 do not disclose the entire LSA-3 sequence, which is the subject of the present application. Moreover, the partial LSA-3 sequence disclosed in these issued patents is furthermore excluded from the claims.

To further evidence that the present invention was not derived from Ms. Claudine Guerin-Marchand, Applicants provide the **enclosed** Rule 132 Declarations executed independently by Pierre Druilhe (the Druilhe Declaration) and Pierre Daubersies (the Daubersies Declaration). In the Druilhe Declaration, the history of the joint invention of Ms. Claudine Guerin-Marchand and Mr Pierre Druilhe appearing in U.S. Patent Nos. 6,319,502 B1 and 6,270,771 is described. Also in the Druilhe Declaration, Mr. Pierre Druilhe describes the role of the inventors of the present application. Further, in the Daubersies Declaration, Mr. Pierre Daubersies describes his role in the invention claimed in the present application. Therefore, Applicants submit that the presently claimed invention was not derived from Ms. Claudine Guerin-Marchand.

In view of the foregoing, withdrawal of this rejection is respectfully requested.

The rejections of (a) Claims 27-29, 35, 39 and 43 under 35 U.S.C. § 102 (b) over Guerin-Marchand et al (WO92/13884) in the light of the translation provided in U.S. Patent No. 6,270,771, and (b) Claims 27-29, 35, 39 and 43 under 35 U.S.C. § 102 (e) over Guerin-Marchand (U.S. Patent No. 6,319,502) are respectfully traversed.

U.S. Patent No. 6,270,771 discloses various sequences from LSA-1, which have the following repetitive sequence:

$X_1$DLEQ$X_2$R$X_3$AKEKLQ$X_4$QQQ

where $X_1$ is Ser or Arg

$X_2$ is Glu or Asp

$X_3$ is Arg or Leu

$X_4$ is Glu or Gly.

This sequence is not present in SEQ ID NO: 3 of the present invention. Thus, even sequences which are longer that include the above sequence are not claimed in the present invention.

The other sequences described in U.S. Patent No. 6,270,771 are derived from the clone DG729S. The 729Sprotein is described at column 5 as SEQ ID No:24 and has 151 amino acids. As pointed out by the Examiner, SEQ ID NO: 11 of the present invention is not disclosed in SEQ ID NO: 3. The remaining sequence of SEQ ID NO: 24 in the prior art is excluded from the presently claimed invention as SEQ ID NO: 10.

The remaining sequences in U.S. Patent 6,270,771, which are described are SEQ ID NO: 25, which is the equivalent sequence to excluded SEQ ID NO: 12 of the present invention; SEQ ID NO:26, which is the equivalent excluded sequence of SEQ ID NO: 13 of the presently claimed invention; SEQ ID NO: 27, which is the equivalent to the excluded SEQ ID NO: 14 of the presently claimed invention; and SEQ ID NO: 28, which is the equivalent to the excluded SEQ ID NO: 15 of the presently claimed invention.

Furthermore, SEQ ID NOS: 25 to 28 have amino acid sequences that are less than 50 consecutive amino acids. As pointed out by the Examiner, U.S. Patent 6,270,771 discloses that the invention encompasses "any sequence comprising 4 to 5 amino acids up to the maximal number of the sequences described above." Since, SEQ ID NOs. 25 to 28 have 47, 26, 27 and 27 total amino acids, these sequences cannot be said to encompass a polypeptide molecule having at least 50 consecutive amino acids of SEQ ID NO: 3 of the present invention.

Therefore, U.S. Patent No. 6,270, 771 fails to disclose or suggest the presently claimed invention. In order for a reference to anticipate an invention, the reference "must teach every element of the claim" (MPEP §2131). Accordingly, U.S. Patent No. 6,270, 771 does not anticipate the invention as presently claimed.

For the same reasons as stated above, the presently claimed invention is not anticipated by U.S. Patent 6,319, 502, which is a divisional of U.S. 6,270,771. Since both issued patents have the same specification, U.S. Patent 6,319, 502 does not anticipate the claimed invention. In summary U.S. Patent 6,319,502 teaches sequences that are not presently being claimed, are excluded from the claims or are not of sufficient length to anticipate the presently claimed invention.

Therefore, in view of the above, withdrawal of these rejections is respectfully requested.

Finally, Applicants remind the Examiner that MPEP §821.04 states:

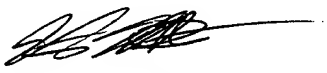
> ...if applicant elects claims directed to the product, and a product claim is subsequently found allowable, withdrawn process claims which depend from or otherwise include all the limitations of the allowable product claim *will be rejoined.* (emphasis added)

Accordingly, should the elected invention be found allowable, Applicants request that withdrawn process claims be rejoined and examined.

Applicants submit that the application is in condition for allowance. Early notice to this effect is earnestly solicited.

Respectfully submitted,

OBLON, SPIVAK, McCLELLAND,
MAIER & NEUSTADT, P.C.
Norman F. Oblon



Vincent K. Shier, Ph.D.
Registration No. 50,552

Customer Number
22850

Tel: (703) 413-3000
Fax: (703) 413-2220
(OSMMN 08/03)

PATENT

Attorney Docket No.200773USODiv

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

| | | |
|---|---|---|
| In re Patent Application of | ) | |
| | ) | |
| Pierre DRUILHE et al | ) | |
| | ) | Group Art Unit: 1641 |
| Serial No.:09/742,096 | ) | |
| | ) | Examiner: James Leslie GRUN |
| Filed: December 22, 2000 | ) | |
| | ) | |

For: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES

**1.132 Declaration**

Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

I, Pierre Daubersies do hereby declare the following:

(1) I Am currently working in the Research and Development Department at Merck LIPHA SANTE in France. Prior to being employed by Merck, I worked in the Laboratory of Parisitology at Institut Pasteur under the head of the Department, Mr. Pierre Druilhe. During the years of 1992 to 1994, I was a post-doctoral Fellow in Mr. Druihle's laboratory working on a

project to clone, sequence and characterize LSA-3. My *Curriculum Vitae* is enclosed. I am one of the inventors of the above-captioned U.S. patent application.

(2) In the present invention the entire *lsa-3* gene was cloned and sequenced, which was the goal of my post-doctorate studies.. To clone and sequence the entire *lsa-3* gene was in fact quite difficult since the genome of *Plasmodium falciparum* is very rich in A: T bases (80%), thus there is a rarity of restriction sites which can be used and also instability of cloning certain fragments when they are inserted into plasmid vectors. Thus, it took further ingenuity to obtain full length LSA-3, analyze its sequence, as well as to determine whether there was any polymorphism in different isolates, which is generally the case between different clones, strains and isolates of *Plasmodium*, as well as to determine the antigenicity in chimpanzees of LSA-3 and fragments thereof.

(3)Secondly, I am one of the true inventors of the claimed subject matter of the present patent application and did not derive the claimed subject matter from Ms. Claudine Guerin-Marchand, who at that epoch was working in the same laboratory. Although Ms. Claudine Guerin-Marchand did obtain a clone 729S while working under the supervision of Professor Druilhe, this was only a 151 amino acid antigen of LSA-3, which is a short fragment.

(4) I further declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of this application or any patent issuing thereon.

13-03-05

---

Date

---

Pierre Daubersies

PATENT

Attorney Docket No.200773USODiv

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

| | | |
|---|---|---|
| In re Patent Application of | ) | |
| | ) | |
| Pierre DRUILHE et al | ) | |
| | ) | Group Art Unit: 1641 |
| Serial No.:09/742,096 | ) | |
| | ) | Examiner: James Leslie GRUN |
| Filed: December 22, 2000 | ) | |
| | ) | |

For: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES

**1.132 Declaration**

Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

I, Pierre Druilhe do hereby declare the following:

(1) I am currently the director of the Biomedical Parasitology Unit at Institut Pasteur in Paris, France and have been the director of this department for many years. As can be seen from my attached *Curriculum Vitae*, I have published 260 articles and have 10 U.S. Patents issued in my name. I am one of the inventors of the above-captioned U.S. patent application.

(2) The primary scientific interest in my laboratory is focused on the understanding of *Plasmodium falciparum*/human immune reactions, both at the level of the pre-erythrocytic stages (sporozoite and liver stage), as well as the merozoite stage. The main activities with these scientific studies in my laboratory are aimed at vaccines to treat or prevent malaria, as well as drug discovery in this area.

(3) As head of a laboratory at Institut Pasteur, I supervise and am responsible for many researchers, scientific trainees, as well as other personnel such as laboratory technicians and office staff. My laboratory also collaborates on particular scientific projects with various scientific groups not associated with Institut Pasteur.

(4) I have read and understood the last U.S. Official Action dated December 3, 2004. As I understand, the Examiner has rejected claims 27 to 29, 35, 39 and 43 on the basis that Mr. Pierre Daubersies and I did not invent the subject matter which is currently claimed and sought to be patented. It is my further understanding that the rationale behind this rejection is that issued U.S. Patents 6,270,771 B1 and 6, 319,502 B1 have myself and Claudine Guerin-Marchand as inventors, while in the above present patent application the other inventor is Pierre Daubersies. Since a partial sequence of a 729S protein was disclosed in issued patents 6,270,771 and 6,319,502, the Examiner maintains that the present inventorship is not correct. I disagree with the Examiner's conclusions for the following reasons.

(5) U.S. Patent Nos. 6,270,771 and 6,319,502 are based on the discovery and sequencing of peptides from LSA-1 (liver stage antigen 1). At that epoch and as described in those patent specifications, one clone that was derived from our extensive search of the genomic library of *Plasmodium falciparum* was directed to a protein different from LSA-1; i.e., LSA-3 clone DG729S. The sequencing of this clone revealed a polypeptide of 151 amino acids, which is only a partial sequence of the LSA-3 protein. This scientific work was in fact the conception of myself

and Ms. Claudine Guerin-Marchand, who was a researcher in my laboratory at that time. After appropriate evaluation by Institut Pasteur's Patent and Invention Office, it was decided that Claude Guerin-Marchand and myself were inventors of at least one claim for U.S. Patent 6,270,771 and 6,319,502.

(6) In contrast, in the present invention the entire *lsa-3* gene was cloned and sequenced. To clone and sequence the entire *lsa-3* gene was in fact quite difficult since the genome of *Plasmodium falciparum* is very rich in A: T bases (80%), thus there is a rarity of restriction sites which can be used and also instability of cloning certain fragments when they are inserted into plasmid vectors. Thus, it took further ingenuity to obtain full length LSA-3, analyze its sequence, as well as to determine whether there was any polymorphism in different isolates, which is generally the case between different clones, strains and isolates of *Plasmodium*, as well as to determine the antigenicity in chimpanzees of LSA-3 and fragments thereof.

(7) In this respect, a post-doctoral fellow, Mr. Pierre Daubersies undertook the project to clone LSA-3 with myself. After appropriate evaluation by Institut Pasteur's Patent and Invention Office, it was concluded that myself and Mr. Duabersies were considered inventors for at least one claim in the present patent application.

(8) I am one of the true inventors of the claimed subject matter of the present patent application and did not derive the invention from Ms. Claudine Guerin-Marchand.

(9) I further declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of this application or any patent issuing thereon.

7-03-2005

Date

Pierre Druilhe

Dept.: CHEMICAL

OSMM&N )No. 200773US0DIV By: NFO:JK:jot

Serial No. NEW APPLICATION

In the matter of the Application of: Pierre DRUILHE et al.

For: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES

The following has been received in the U.S. Patent Office on the date stamped hereon:

- ■ 41 pp. Specification & 26 Claims/Drawings 26 Sheets
- ■ Copy of Combined Declaration, Petition & Power of Attorney 3 pages (executed)
- □ List of Inventor Names and Addresses
- ■ Utility Patent Application
- ■ Notice of Priority w/PCT/IB/304
- ■ Check for $854.00
- ■ Fee Transmittal Form
- □ Assignment/PTO 1595 pages:
- □ Letter to Official Draftsman
- □ Letter Requesting Approval of Drawing Changes
- □ Drawings sheets □ Formal
- □ Letter
- ■ Preliminary Amendment
- ■ Information Disclosure Statement
- ■ Cited References (6)
- ■ International Search Report
- □ Statement of Relevancy
- □ IDS/Related/List of Related Cases
- □ Restriction Response
- □ Rule 132 Declaration
- □ Petition for Extension of Time
- □ Notice of Appeal
- □ Brief
- □ Issue Fee Transmittal
- ■ White Advance Serial Number Card
- □ Small Entity Status is Claimed
- ■ Substitute Sequence Listing w/Attached Sequence Alignment
- □



- □ CPA
- □ Priority Doc
- ■ Dep. Acct. Order Form



- ■ PTO-1449
- □ Cited Pending Applications
- □ Election Response

Due Date: December 26, 2000

[redacted]

Dept.: Chemical

OSMM&N File No. 200773US0DIV

By: NFO/DJP/km

Serial No. 09/742,096

In the matter of the Application of: Pierre DRUILHE, et al.

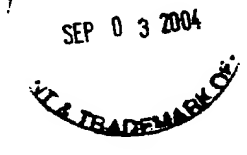
For: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES

**Due Date: September 3, 2004**

The following has been received in the U.S. Patent Office on the date stamped hereon

- Credit Card Form for $950.00
- Dep. Acct. Order Form
- PTO Cover Letter
- Amendment and Request for Reconsideration w/Annex I attached
- Petition for Extension of Time (Three-Months)
- Sequence Listing (Paper)
- Computer-Readable Sequence Listing (Diskette)
- Certified English Translation of French Application No. 9507007
- Letter Submitting Replacement Drawing Sheets for FIGS. 1a-1f, 2a-21, 3a-3d 4b, 5, 6, 7, 8, 9a-9c, 10a-10b (attached)

COPY



Docket No. 200773US0DIV

# IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN RE APPLICATION OF: Pierre DRUILHE et al. GAU:

SERIAL NO: NEW APPLICATION EXAMINER:

FILED: HEREWITH

FOR: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES



## REQUEST FOR PRIORITY

ASSISTANT COMMISSIONER FOR PATENTS
WASHINGTON, D.C. 20231

SIR:

■ Full benefit of the filing date of U.S. Application Serial Number 08/973,462, filed February 6, 1998, is claimed pursuant to the provisions of **35 U.S.C. §120**.

□ Full benefit of the filing date of U.S. Provisional Application Serial Number , filed , is claimed pursuant to the provisions of **35 U.S.C. §119(e)**.

■ Applicants claim any right to priority from any earlier filed applications to which they may be entitled pursuant to the provisions of **35 U.S.C. §119 and §120**, as noted below.

In the matter of the above-identified application for patent, notice is hereby given that the applicants claim as priority:

| COUNTRY | APPLICATION NUMBER | MONTH/DAY/YEAR |
|---|---|---|
| **France** | **95/07007** | **June 13, 1995** |
| **WIPO** | **PCT/FR96/00894** | **June 12, 1996** |

Certified copies of the corresponding Convention Application(s)

□ are submitted herewith

□ will be submitted prior to payment of the Final Fee

□ were filed in prior application Serial No. filed

■ were submitted to the International Bureau in PCT Application Number PCT/FR96/00894.
Receipt of the certified copies by the International Bureau in a timely manner under PCT Rule 17.1(a) has been acknowledged as evidenced by the attached PCT/IB/304.

□ (A) Application Serial No.(s) were filed in prior application Serial No. filed ; and

(B) Application Serial No.(s)

□ are submitted herewith

□ will be submitted prior to payment of the Final Fee

Respectfully Submitted,



OBLON, SPIVAK, McCLELLAND,
MAIER & NEUSTADT, P.C.

Norman F. Oblon
Registration No. 24,618

James J. Kelly, Ph.D.
Registration No. 41,504

22850
Tel. (703) 413-3000
Fax. (703) 413-2220
(OSMMN 10/98)

# TRAITÉ DE COOPÉRATION EN MATIÈRE DE BREVETS

Expéditeur : le BUREAU INTERNATIONAL

## PCT

### NOTIFICATION RELATIVE A LA PRESENTATION DU DOCUMENT DE PRIORITE

(instruction administrative 411 du PCT)

Destinataire:

GUTMANN, Ernest
Yves Plasseraud S.A.
3, rue Chauveau-Lagarde
F-75008 Paris
FRANCE

| | |
|---|---|
| **Date d'expédition** (jour/mois/année)<br>24 juillet 1996 (24.07.96) | |
| **Référence du dossier du déposant ou du mandataire**<br>B2775A-FL | **NOTIFICATION IMPORTANTE** |

| Demande internationale no | Date du dépôt international | Date de priorité |
|---|---|---|
| PCT/FR96/00894 | 12 juin 1996 (12.06.96) | 13 juin 1995 (13.06.95) |

**Déposant**

INSTITUT PASTEUR etc

La date de réception par le Bureau international du ou des documents de priorité correspondant à la ou aux demandes suivantes est notifiée au déposant:

| Demande antérieure no: | Date de priorité: | Pays dans lequel ou pour lequel la demande a été déposée: | Date de réception du document de priorité |
|---|---|---|---|
| 95/07007 | 13 jui 1995 (13.06.95) | FR | 23 jui 1996 (23.07.96) |

COPY

**Bureau international de l'OMPI**
**34, chemin des Colombettes**
**1211 Genève 20, Suisse**

no de télécopieur: (41-22) 740.14.35

Fonctionnaire autorisé:

G. Bähr

no de téléphone: (41-22) 730.91.11



001165884

COPY

# VERIFICATION OF A TRANSLATION

I, the below named translator, hereby declare that:

My name and post office address are as stated below:

That I am knowledgeable in the English language and in the language in which the below identified French Application was filed, and that I believe the English translation of the French Application N° 9507007 filed on 13/06/95 in the name of Institut Pasteur

is a true and complete translation of the above identified French Application as filed.

Date: ....28/04/2000....

Full name of the translator — Véronique MARCADÉ

Signature of the translator

Post Office Address — 3, rue Chauveau-Lagarde
F-75008 Paris
FRANCE

# POLYPEPTIDE MOLECULES OF THE PRE-ERYTHROCYTIC STAGE OF MALARIA

The parasites responsible for malaria in man display different morphologies in the human host and express different antigens depending on their location in the body. The morphological and antigenic differences of these parasites during their life cycles in man enable different stages of development in the liver and in the blood to be defined: the sporozoite, the infectious form injected by the vector mosquito, transforms rapidly into a schizont in the host's hepatocytes and thereafter infects the erythrocytes. The intrahepatic localization of P.falciparum manifests itself in the expression of a group of antigens specific to this stage of development and which are highly immunogenic under the natural conditions of exposure to the disease. This clinically silent phase is at present the only one against which a very strong, sterilizing immunity can be induced experimentally in man, by injecting irradiated sporozoites capable of entering the hepatocyte and of developing therein but without being able to lead on to the blood stage of the disease. Accordingly, the inventors have concentrated the bulk of their efforts on these two pre-erythrocytic stages. However, these stages are also the most intricate ones to study, and hence the least understood, since it is difficult or even impossible to obtain biological material, the only in vitro study

model affords a very low yield and the best animal model remains the chimpanzee, the use of which is limited and expensive.

In order to gain access to the antigens of the pre-erythrocytic stages, the inventors used sera of individuals who had resided for 25 years in a region where the disease is endemic but who were on permanent prophylaxis with chloroquine. These individuals were regularly subjected to infected mosquito bites but did not develop any complete blood infection. Their serum hence contained antibodies directed essentially against the pre-erythrocytic stages, which was verified by immunofluorescence (IF) and western blotting on the 3 stages of the parasite.

The use of these sera for screening a library of genomic DNA of the parasitic clone of P.falciparum, the library being constructed in expression vectors in a phage lambda gt11 (V. Rosario, Science 212, 1981, pp. 1037-1038; and Thaithong et al., Transactions of Royal Society of Tropical Medicine and Hygiene, 1984, 78:242-245), led to the demonstration of polypeptides of the pre-erythrocytic stage, in particular the SALSA (sporozoite liver stage antigen) polypeptides described in EP A-0,407,230 and LSA-1 (liver stage antigen) described in WO 92/13884. The present invention relates to new polypeptide molecules specific to the pre-erythrocytic stage, and to their use as active principle of antimalarial vaccine or in methods of diagnosis of the disease.

The invention is the outcome of the demonstration by the inventors of the special properties of a particular antigen referred to LSA-3 and of its fragments, which are seen to be candidates with a strong potential for producing an antimalarial vaccine, for the following reasons:

a) when a fraction of LSA-3 was used in combination with another antigen of the same stage of development of the parasite, such as LSA-1, to immunize

chimpanzees, the animal responding to both molecules or only to LSA-3 displays the feature of not having parasites in the blood, of having a substantial decrease of the parasites in the liver and of manifesting a substantial recruitment of mononuclear cells indicating a response in terms of cellular immunity;

b) in regions where the disease is endemic, a very clear correlation is observed between the protection of individuals against natural infection by sporozoites and their responses in terms of antibodies against LSA-3;

c) in eight human volunteers immunized by injection of irradiated sporozoites, antibodies directed against LSA-3 are found in each of the four individuals resisting sporozoite infection and in none of the other four volunteers who developed a blood infection;

d) antibodies obtained against the peptide DG729 in WO 92/13884, already described, give a cross-reaction with the sporozoite and liver stages of the murine parasite P.yoelii, which permits a significant exploitation of the mouse model. In vitro, the human antibodies immunopurified on DG729 are capable, even at very low concentrations, of blocking the entry of P.yoelii sporozoites into mouse hepatocytes. In vivo, mice immunized with DG729 are fully or partially protected against infection by P.yoelii sporozoites;

e) lastly, some epitopes, in particular in the non-repeat portions of the molecule, stimulate the secretion of interferon-γ by monocytes, this mediator enabling the intrahepatic development of the parasite to be inhibited (S. Mellouk et al., The Jour. Of Immun. 139, 4192-4195, 1987);

All these properties, some of which will be demonstrated in detail in the experiments described later, show that the LSA-3 antigen displays both good antigenicity and good immunogenicity.

The inventors were able to confirm and define the specificity of the stages of expression of the

molecule; in the sporozoites, this expression was confirmed by the surface immunofluorescence of several strains and isolates. In Western Blot analysis, the LSA-3 molecule appears as a protein of molecular weight 200,000 daltons. While the messenger RNAs of sporozoites could not be obtained in sufficient amounts for a northern blot analysis, reverse PCR experiments confirmed the expression of LSA-3 at this stage. In infected hepatocytes, LSA-3 is observed in the parasitophorous vacuole of the parasite by immunofluorescence using antibodies against the repeat and non-repeat regions of the protein, as well as by electron microscopy.

A fragment of LSA-3 designated 729S, as well as three peptides designated NRI and NRII included in the non-repeat portion and 729R included in the repeat portion, have been described in Application WO 92/13884. Nevertheless, this document does not mention the special properties mentioned above, or other fragments of LSA-3 which could be either longer or shorter, included or combined with these fragments, which might display especially advantageous properties for use in vaccines.

The subject of the invention is polypeptide molecules containing at least ten consecutive amino acids of the amino acid sequence shown in Figure 2 and designated SEQ ID No. 2, and representing LSA-3, the following polypeptides being excluded:

- RDELFNELLNSVDVNGEVKENILEESQVNDDIFNSLVKSVQQEQQHNVEE
- VEESVEENDEESVEENVEENVENNDDGSVASSVEESIASSVDESIDSSIE-ENVAPTVEEIVAPTVEEIVAPSVVEKCAPSVEESVAPSVEESVAEMLKER (729S)
- RDELFNELLNSVDVNGEVKENILEESQVNDDIFNSLVKSVQQEQQHN
- DELFNELLNSVDVNGEVKENILEESQ, (NRI)
- LEESQVNDDIFSNSLVKSVQQEQQHNV, (NRII)
- VESVAPSVEESVAPSVEESVAENVEESV. (729RE)

Other molecules according to the invention contain at least 20 consecutive amino acids or at least 50.

This set of polypeptides and the LSA-3 molecule are, throughout hereinafter, "polypeptides of the invention".

The experimental results and the comparisons of sequences between different P.falciparum isolates indicate the existence of at least 70% homology between equivalent antigens of the liver stage of the parasite. Thus any peptide molecule displaying at least 70% homology with any one of the molecules defined above forms part of the invention, as do those displaying at least 70% homology with the following sequence:

Leu Leu Ser Asn Ile Glu Glu Pro Lys Glu Asn Ile Ile Asp Asn Leu Leu Asn Asn Ile (CT1)

lying between amino acids 140 and 159 of K1 or 23 and 42 of T9/96.

Likewise forming part of the invention are the polypeptide molecules displaying at least 70% homology with the sequence depicted in Figure 3, which depicts a portion of LSA-3 in T9/96: the DNA of this P.falciparum isolate was digested with restriction enzymes, then cloned into lambda gt11 and thus enabled the gene library of this isolate, already described above, to be constituted.

Conjugates consisting of a polypeptide originating from LSA-3 linked covalently via a lysine bridge to saturated or unsaturated lipid residues also form part of the invention, more especially when the lipid residue is a palmitoyl or a palmityl or an oleyl. $C_{16}$ or $C_{18}$ residues were thus coupled via a lysine bridge to the peptides NRI, NRII, 729RE and CT1 already depicted above. The method of synthesis used for these conjugates is described in Bourgault, Journal of Immunology, 149, 3416 (1992) and Rouaix, Vaccine, 12, 1209 (1994).

The invention also covers immunogenic compositions containing at least one polypeptide molecule or one conjugate described above, as well as the vaccines containing these immunogenic compositions. Other immunogenic epitopes, in particular LSA-1, SALSA, STARP, have already been described in EP A-0407230 and in WO 92/13884. The vaccine compositions according to the invention can advantageously contain a mixture of immunogenic peptides originating from LSA-3 and of the peptides or antigens originating from LSA-1, SALSA or STARP; a more especially advantageous mixture could be the one consisting, on the one hand of NRI, NRII or whole LSA-3, these being coupled or otherwise to a lipid residue, and on the other hand the peptides SALSA-1, SALSA-2 or the SALSA antigen coupled or otherwise to a lipid residue.

All polypeptide molecules corresponding to the above definition and displaying at least 70% homology with the polypeptides LSA-3, CT1, NRI, NRII or 729RE may be combined in homologous or heterologous fashion with other peptide sequences or sequences originating from another antigen of the different stages of P.falciparum.

The invention also covers the polyclonal or monoclonal antibodies which specifically recognize the polypeptide molecules of the invention.

These molecules of the invention may be used for carrying out diagnostic methods and producing kits enabling the existence of P.falciparum infection to be detected; this method can be either an assay of circulating specific antibodies, by carrying out standard serological methods by bringing one of the above antigens into contact with a biological fluid of the individual in question, or methods of assay of antigens using polyclonal or monoclonal antibodies obtained by standard methods for obtaining such antibodies with the corresponding antigens. In the diagnostic outfits or kits of the invention, the

reagents enabling the antigen/antibody complexes produced to be detected, which can also carry a label or be capable of being recognized in their turn by a labelled reagent, are present. Depending on whether it is desired to carry out an antigen test or a serological test, the kit comprises either the antibodies or the antigens of the invention.

The invention also covers all the nucleotide sequences coding for a polypeptide of the invention, as well as any recombinant nucleic acid containing at least one nucleotide sequence of the invention, inserted into a nucleic acid which is heterologous with respect to the said nucleotide sequence.

The nucleic acid sequences coding for LSA-3 or its immunogenic fragments and corresponding to one of the following definitions form part of the invention:

(a) the linked succession of nucleotides as depicted in SEQ ID No. 1 of Figure 1, or

(b) the linked succession of nucleotides depicted in SEQ ID No. 2 of Figure 2,

(c) a linked succession displaying at least 70% homology with that of Figure 1 or of Figure 2, or

(d) a linked succession of nucleotides which are complementary to those presented in (a), (b) or (c).

The expression "coding for LSA-3" is understood to refer both to the gene depicted in SEQ ID No. 1 of Figure 1 and the cDNA depicted in SEQ ID No. 2 of Figure 2.

The invention relates more especially to a recombinant nucleic acid in which the nucleotide sequence of the invention is preceded by a promoter (in particular an inducible promoter), under the control of which the transcription of the said sequence is capable of being performed, and, where appropriate, followed by a sequence coding for transcription termination signals.

The invention also covers the coding sequence originating from the clone T9/96 depicted in Figure 3 by SEQ ID No. 3.

In this sequence, the fragment CT1 lies between nucleotides 67 and 126, the fragment 679 begins at nucleotide 206 and the fragment 729RE lies between nucleotides 547 and 630.

Lastly, the invention covers any recombinant vector used especially for the cloning of a nucleotide sequence of the invention, and/or for the expression of the polypeptide encoded by this sequence, and characterized in that it contains a recombinant nucleic acid as defined above in one of its sites which is not essential for its replication.

As an example of an abovementioned vector, plasmids, cosmids, phages or viruses may be mentioned.

As such, the invention relates more especially to the plasmid pK 1.2. deposited at the CNCM under the No. I-1573.

The subject of the invention is also a method for preparing a polypeptide of the invention, by transformation of a cell host using a recombinant vector of the abovementioned type, followed by the culturing of the cell host thus transformed and the recovery of the polypeptide in the culture medium.

Thus, the invention relates to any cell host transformed by a recombinant vector as defined above, and comprising the regulatory elements permitting the expression of the nucleotide sequence coding for a polypeptide according to the invention.

The invention likewise covers DNA (or RNA) primers which can be used in the context of the synthesis of nucleotide and/or polypeptide sequences of the invention, by the PCR (polymerase chain reaction) technique or any other method known at the present time for amplifying nucleic acids, such as LCR, CPR, ERA, SPA, NASBA, and the like.

The invention relates to any DNA or RNA primer, characterized in that it consists of approximately 10 to 25 nucleotides which are identical or complementary to the first 10 to 25 nucleotides of the nucleotide sequence coding for a peptide sequence according to the invention, or identical to the last 10 to 25 nucleotides of the said sequence.

Thus, the present invention also covers a method for preparing a polypeptide of the invention comprising the following steps:

- where appropriate, the prior amplification by standard techniques of the amount of nucleotide sequences coding for the said polypeptide using two suitably chosen DNA primers,
- the culturing, in a suitable culture medium, of a cell host previously transformed by a vector containing a nucleic acid according to the invention comprising the nucleotide sequence coding for the said polypeptide, and
- the recovery from the abovementioned culture medium of the polypeptide produced by the said transformed cell host.

By way of example of DNA or RNA primers according to the invention, the following pairs of sequences may be mentioned:

S1: GTGATGAACTTTTTAATGAATTATTAAA (SEQ ID No. 4)
S2: TGTTGTTCTTGTTGAACACTTTTTACTAA (SEQ ID No. 5)

whose respective positions on the LSA-3/K1 gene depicted in Figure 1 are from 695 to 722 and from 829 to 799 (reading in the reverse direction), or the pair:

6.1: GGTATCGAAACTGAGGAAATAAAGG (SEQ ID No. 6)
6.2: CATAGCAGGAACATCAACATCCAC (SEQ ID No. 7)

whose respective positions are 2668 to 2692 for 6.1 and 3456 to 3433 for 6.2 (reading in the reverse direction).

The peptides of the invention may also be prepared by the standard techniques of peptide synthesis. This synthesis may be carried out in

homogeneous solution or in the solid phase. For example, use may be made of the technique of synthesis in homogeneous solution described by Houben-Weyl in the work entitled "Methoden der Organischen Chemie" (Methods in Organic Chemistry) edited by E. Wunsch, vol. 15-I and II. Thieme, Stuttgart 1974, or that described by R.D. Merrifield in the paper entitled "Solid phase peptide synthesis" (J. Am. Chem. Soc., 45, 2149-2154).

The invention also covers the water-soluble oligomers of the abovementioned monomeric peptides.

Oligomerization can cause an enhancement of the immunogenicity of the monomeric peptides according to the invention. While such numerical information cannot be regarded as limiting, it may nevertheless be mentioned that these oligomers can, for example, contain from 2 to 10 monomer units.

To carry out the oligomerization, use may be made of any polymerization technique commonly used in the peptide field, this polymerization being conducted until an oligomer or polymer containing the requisite number of monomer motifs for acquiring the desired immunogenicity is obtained.

One method of oligomerization or polymerization of the monomer consists in reacting the latter with a crosslinking agent such as glutaraldehyde.

Use may also be made of other oligomerization or coupling methods, for example the one employing successive couplings of monomer units via their carboxy- and amino-terminal functions in the presence of homo- or heterobifunctional coupling agents.

The invention also relates to the conjugates obtained by covalent coupling of the peptides according to the invention (or of the abovementioned oligomers) to physiologically acceptable and non-toxic (natural or synthetic) carrier molecules that enable, in particular, the immunogenicity to be increased, via complementary reactive groups carried, respectively, by

the carrier molecule and the peptide. By way of example of macromolecular carrier molecules or supports which participate in the constitution of the conjugates according to the invention, there may be mentioned natural proteins such as tetanus toxoid, ovalbumin, serum albumins, haemocyanins, tuberculin PPD (PPD: purified protein derivative), and the like.

By way of synthetic macromolecular supports, there may be mentioned, for example, polylysines or poly(DL-alanine)-poly(L-lysine)s.

By way of hydrocarbon or lipid supports, there may be mentioned saturated or unsaturated fatty acids, and preferably $C_{16}$ or $C_{18}$ acids of the oleic or palmitoleic type.

To synthesize the conjugates according to the invention, use may be made of methods which are known per se, such as the one described by Frantz and Robertson in Infect. and Immunity, 33, 193-198 (1981), or the one described in Applied and Environmental Microbiology (October 1981), vol. 42, No. 4, 611-614 by P.E. Kauffman, using the peptide and the appropriate carrier molecule.

The nucleic acids of the invention may be prepared either by a chemical method or by other methods.

A suitable method of preparing the nucleic acids of the invention containing not more than 200 nucleotides (or 200 bp in the case of double-stranded nucleic acids) comprises the following steps:

- DNA synthesis using the automated β-cyanoethyl-phosphoramidite method described in Bioorganic Chemistry 4; 274-325 (1986),
- cloning of the nucleic acids thereby obtained into a suitable vector and recovery of the nucleic acid by hybridization with a suitable probe.

A chemical method of preparation of nucleic acids of length greater than 200 nucleotides has already been described in WO 92/13884.

The invention also relates to diagnostic kits which contain one or more amplification primers specific for the LSA-3 gene and which enable the presence of the gene or of the mRNA to be detected in an individual likely to be infected by P.falciparum.

The invention also covers pharmaceutical or vaccine compositions in which at least one of the products according to the invention is present in combination with solid or liquid, pharmaceutically acceptable excipients suitable for the construction of oral, ocular or nasal dosage forms, or excipients suitable for the construction of dosage forms for rectal administration, or alternatively with gelatinous excipients for vaginal administration. It also relates to isotonic liquid compositions containing at least one of the conjugates according to the invention, suitable for administration to the mucosae, in particular the ocular or nasal or pulmonary mucosae.

Advantageously, the vaccine compositions according to the invention contain, in addition, a vehicle such as polyvinylpyrrolidone which facilitates the administration of the vaccine. In place of polyvinylpyrrolidone, it is possible to use any other type of adjuvant, in the traditional sense which was formerly given to this expression, that is to say a substance which enables a medicinal product to be absorbed more readily or which facilitates its action in the body. By way of examples of other adjuvants of this latter type, there may also be mentioned carboxymethylcellulose, aluminium hydroxides and phosphates, saponin or all other adjuvants of this type which are well known to a person skilled in the art. Lastly, they contain, if necessary, an immunological adjuvant, in particular of the muramyl peptide type.

The invention also relates to pharmaceutical compositions containing as active substance at least one of the polyclonal or monoclonal antibodies defined

above, in combination with a pharmaceutically acceptable vehicle.

Additional features of the invention will also become apparent in the examples illustrated with the figures which follow, and show the special features of the molecules of the invention relative to other antigens of the pre-erythrocytic stage of the parasite.

Figure 1 depicts the genomic DNA sequence ID No. 1 of 6152 base pairs of the LSA-3 gene; it originates from the clone K1.2, which itself originates from a Thai isolate.

Figure 2 depicts the cDNA sequence ID No. 2 and the polypeptide sequence of the LSA-3 antigen. The DNA sequence represents 5361 base pairs.

Figure 3 depicts the sequence ID No. 3 of the portion sequenced in the parasite clone T9/96 (1890 base pairs), the upper line being the nucleotide sequence and the lower line the peptide sequence. In this clone, the CT1 sequence lies between nucleotides 67 and 126, the actual fragment DG679 beginning at nucleotide 207. The fragment 729RE lies between nucleotides 547 and 629.

Figure 4a depicts diagrammatically the relative positions of the repeat and non-repeat sequences, the introns and the exons in strains K1 and T9/96, the clones 679 and 729 originating from the latter.

Figure 4b depicts the HCP (hydrophobic cluster plot) of the peptide sequence of the clone DG729.

Figure 5 depicts the amounts of immunoglobulins produced in the serum of chimpanzee Nuria before and after immunization with different LSA-3 peptides.

Figure 6 shows the specific antibody titre of different species of mice immunized either with a peptide or with a corresponding lipopeptide.

Figure 7 shows the inhibition of the sporozoite invasion of liver cells by hyperimmune sera obtained after immunization with different peptides and immunopurified against whole LSA-3.

Figure 8 depicts the comparison of an antigen originating from LSA-3 with two other antigens with respect to type T immunity.

Figure 9 depicts the induction of interferon-γ in the chimpanzees Gerda and Dirk with the peptides originating from the LSA-3 molecule.

Example 1: Cloning and sequencing of the LSA-3 gene

1) Sequencing

Initial screening of the gene library originating from the parasite clone T9/96 with the serum of a missionary treated continuously by prophylaxis enabled us to isolate 120 clones corresponding to molecules expressed at the sporozoite and/or liver stage of the P.falciparum cycle. The clone 729S was used as probe to screen a genomic library of the Thai strain K1 already mentioned above, which contains large EcoR I fragments cloned into phage lambda gt10. A 6.85-kilobase insert containing the whole gene was purified from this gene library and recloned into a pUC18 plasmid for sequencing and characterization. In P.falciparum, the genome of which is very rich in bases A:T(80%), this approach is often rendered difficult by the rarity of restriction sites which can be used, and by the instability or even the impossibility of cloning certain fragments when they are inserted into plasmid vectors.

The structure of the gene is depicted in Figure 4 and displays the following features:

a) a mini-exon 1 coding at its 3′ end for a hydrophobic signal peptide;

b) a short intron (168 base pairs) included between consensus splicing donor and acceptor sites;

c) a second exon of five kilobases which codes for an organized region of 1.8 kilobases, and composed of an arrangement of 7 blocks of 4 amino acids and a 3′ hydrophobic region which might correspond to an

anchorage of the glycosylphosphatidylinositol (GPI) type.

A detailed investigation of the polymorphism of LSA-3 was carried out by sequencing the clone 679, which contains the bulk of the repeat sequences of the LSA-3 gene and a 1-kilobase portion of the 3' non-repeat fraction, the sequence of this fragment being depicted in Figure 3 between nucleotides 207 and 1890.

Comparison of the sequences of the clone 679 originating from P.falciparum clone T9/96, and of the corresponding sequence of LSA-3 originating from the isolate K1, shows that the gene is well conserved, the most significant differences being observed in the repeat region where the blocks of 4 amino acids are well conserved but vary in their number and organization.

In contrast, the non-repeat 5' and 3' portions appear to be especially well conserved, showing up to 100% homology in the 5' region where B and T epitopes have already been identified.

DNA amplifications, in particular by PCR of different P.falciparum strains with 8 primer pairs distributed over the whole of the LSA-3 gene, showed that, except with the ones surrounding the repeat regions, the whole of the genome gives PCR products of similar size, suggesting that the LSA-3 antigen is well conserved.

Various LSA-3 probes, chosen in the repeat and non-repeat regions, were hybridized at low stringency with the DNAs of different species of Plasmodium, and did not enable any gene homologous to LSA-3 to be identified except in the chimpanzee parasite P.reichenowi, confirming the close kinship of this species with P.falciparum.

Surprisingly, the antigen analogous to LSA-3 found in P.yoelii, which gives clear immunological cross-reactions at the surface of the sporozoite with antibodies against the fragment 729S, does not appear

to be conserved at the level of the nucleotide sequence. Lastly, comparison of the LSA-3 sequences with the data bases did not reveal any homology with known molecules, except for the repeat region, some of the motifs of which display a strong analogy with the repeats of a Staphylococcus xylosis gene, but also with two P.falciparum antigens, RESA and Pf11.1, which are both expressed during the blood stage of the parasite. This homology is essentially due to the large amount of "Glu-Glu" sequences in these antigens and in the repeats of LSA-3.

2) Cloning

The insert DG729 and other regions of exon 2 of the strain K1 were cloned into a prokaryotic expression vector pGEX, a vector marketed by the company InVitrogen Corp (San Diego USA). This vector produces a fusion protein with the Schistosoma mansoni glutathione S-transferase (GST), and enables the recombinant proteins to be purified readily by affinity for glutathione-agarose beads. The expression peptides from these vectors are designated:

- for the whole LSA-3 protein: REC protein,
- or for the fragment 729S:729PGEX.

Attempts at cloning other fragments, in particular the fragment 1-5 3NSREP, 3NFREP, 5NR and 5SNREP, caused difficulties related either to the cloning or to the production and purification of the proteins in sufficient amounts for immunization experiments.

Only the fragments 729, NN and 3PC enabled corresponding recombinant polypeptides to be produced and purified in sufficient amounts for analysis of the antigenicity of the molecule.

Example 2: Comparison of the antibody responses of chimpanzee Nuria before and after immunization with different peptides

Figure 5 depicts the amounts of immunoglobulins present in the serum of chimpanzee Nuria before and after immunization with the peptides 729NR1 and 729RE, and the lipopeptides 729NR2 and CT1.

This experiment shows the superiority as regards B immunity of the R antigen, most particularly when it is conjugated to a lipid residue.

Figure 6 shows that the level of specific antibodies measured by ELISA against the peptide 729NR2 in mice immunized with either the peptide 729NRII or the lipopeptide 729NRII is markedly higher when the lipopeptide is used, irrespective of the species of mouse.

Example 3: Effects of the antibodies against the LSA-3 peptides on the inhibition of the entry of sporozoites in mice

The techniques used to prepare the primary hepatocyte cultures, the sporozoites, the antibodies and the indirect fluorescence test are described in detail by S. Mellouk et al., Bulletin of the World Health Organization, 68: 52-59, 1990. The table below compares the results obtained in immunofluorescence, either with antibodies against the fragment 679 or with antibodies obtained against fragments originating from other peptides. The left-hand column shows the number of schizonts detected after 48 h of culture in hepatocytes of Balb/c mice infected by P.yoelii and the right-hand column the same parameters after infection by P.berghei.

| Antibody clones | P.yoelii | | | P.berghei | | |
|---|---|---|---|---|---|---|
| | IFA | No. of LS at 48 h | | IFA | No. of LS at 48 h | |
| | | a) | b) | | | |
| Control | | 88 | 110 | | 119 | 108 |
| 679 | ++ | | 0 | - | | 47 |
| | ++ | | 0 | - | | ND |
| 679 | ++ | 1 | | - | 105 | |
| 679b | ++ | 1 | | - | 133 | |
| 679c | ++ | 1 | | - | 30 | |
| 32 | ++ | 8 | | ± | 103 | |
| 222 | + | | 5 | ± | | 26 |
| 667 | ++ | 276 | 143 | ND | 502 | |
| 362 | + | 3 | | | | |
| 493 | ++ | 55 | | ND | 508 | |
| α P.b. CSP Mab | | | 82 | +++ | | 30 |
| α P.y. CSP Mab | +++ | | 171 | | 138 | |

It is clearly apparent that the antibody against the peptide 679 has an almost complete inhibitory effect on the number of what they [sic] observed at 48 h in the liver cells. Likewise, Figure 7 shows the inhibition of the sporozoite invasion of liver cells by hyper human [sic] sera obtained after immunization with different peptides and immunopurified against whole LSA-3.

Example 4: Cytotoxicity test against the peptide 729NRII in the chimpanzee Gerda

The chimpanzee Gerda was immunized via the i.v. route with the lipopeptide 729NRII originating from the LSA-3 antigen. Blood is drawn 9 days after the 4th injection. The PBMCs were incubated in vitro with 5 µg/ml of the peptide 729NRII (addition of recombinant IL-2, 10 U/ml, on day 3). On day 15, the cytotoxic activity was studied against autologous blasts generated with PHA at a concentration of 0.5 µg/ml. The blasts were preincubated overnight with 5 µg/ml of the peptide 729NRII, and with a control peptide, namely RESA, or without a peptide. The peptides are not added

during the test (8 hours). The number of targets per well is 5000.

PBMCs from Gerda incubated for the same period with 5 µg/ml of a control peptide or the peptide 729NRI (originating from the same antigen) do not bring about the lysis of autologous blasts preincubated or otherwise with the above peptides.

Figure 8 shows the results obtained for an E/T (effector to target) ratio varying from 12 to 0.03. It is seen that the target cells presensitized with the peptide 729NRII are lysed in the presence of effector cells, indicating a cytotoxic T type immune response specific to this antigen.

The lipopeptide NRII injected via the i.v. route is capable, without adjuvant, of inducing a specific cytotoxic response.

Example 5: Effect of the peptide NRI on interferon-γ production

Interferons have been shown to have an inhibitory activity in the development of P.falciparum in human hepatocytes in culture (Sylvie Mellouk et al., The Journal of Immunology, vol. 139 No. 12: 41-92, 41-95, 1987). The results obtained with the peptides of the invention are as follows:

The chimpanzee Gerda, immunized with the polypeptide NR2 and boosted with the recombinant DG729, carries PBMCs capable of secreting high levels of IFN-γ in the presence of the LSA-3 peptides, especially the peptide 729NRI. The result was confirmed in the chimpanzee Dirk, immunized with the same protein. The chimpanzee BRAM, an unimmunized control, does not show any interferon in the blood against the LSA-3 peptides.

The sequences ID n° 1 to 3 are depicted in figures 1 to 3.

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleotide
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GTGATGAACT TTTTAATGAA TTATTAAA 28

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleotide
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TGTTGTTCTT GTTGAACACT TTTTACTAA 29

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleotide
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GGTATCGAAA CTGAGGAAAT AAAGG 25

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleotide
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

CATAGCAGGA ACATCAACAT CCAC 24

## CLAIMS

1. Polypeptide molecules containing at least 10 consecutive amino acids of the amino acid sequence shown in Figure 2, the following polypeptides being excluded:

- RDELFNELLNSVDVNGEVKENILEESQVNDDIFNSLVKSVQQEQQHNVEE
- VEESVEENDEESVEENVEENVENNDDGSVASSVEESIASSVDESIDSSIE-ENVAPTVEEIVAPTVEEIVAPSVVEKCAPSVEESVAPSVEESVAEMLKER (729S)
- RDELFNELLNSVDVNGEVKENILEESQVNDDIFNSLVKSVQQEQQHN
- DELFNELLNSVDVNGEVKENILEESQ, (NRI)
- LEESQVNDDIFSNSLVKSVQQEQQHNV, (NRII)
- VESVAPSVEESVAPSVEESVAENVESSV. (729RE)

2. Molecules according to Claim 1, characterized in that they contain at least 20 consecutive amino acids of the said sequence.

3. Molecules according to Claim 2, characterized in that they contain at least 50 consecutive amino acids of the said sequence.

4. Polypeptide molecule displaying at least 70% homology with one of the molecules of any one of Claims 1 to 3.

5. Polypeptide molecule, characterized in that it displays at least 70% homology with the following sequence:
Leu Leu Ser Asn Ile Glu Glu Pro Lys Glu Asn Ile Ile Asp Asn Leu Leu Asn Asn Ile (CT1).

6. Polypeptide molecule according to one of Claims 1 to 4, characterized in that it displays at least 70% homology with the sequence depicted in Figure 3.

7. Immunogenic composition, characterized in that it contains at least one polypeptide molecule according to any one of Claims 1 to 6 and at least one pharmaceutical vehicle.

8. Antimalarial vaccine composition containing, among other immunogenic principles, a polypeptide molecule according to one of Claims 1 to 6.

9. Vaccine composition according to Claim 8, characterized in that it contains, in addition, a molecule containing at least one epitope and which originates from the group consisting of the LSA-1, SALSA or STARP molecules.

10. Composition according to Claim 9, characterized in that it contains at least two immunogens, the first being chosen from the following polypeptides:

- that of Figure 2,
- NRI,
- NRII,

and the second being chosen from the group consisting of SALSA, SALSA I and SALSA II.

11. Polyclonal or monoclonal antibodies which specifically recognize the polypeptide molecules according to any one of Claims 1 to 6.

12. Method of in vitro diagnosis of malaria in an individual likely to be infected by P. falciparum, which comprises the bringing of a tissue or biological fluid taken from an individual into contact with a molecule according to one of Claims 1 to 8, under conditions permitting an immunological reaction, said polypeptide molecule and antibodies possibly present in the tissue or the biological fluid, and the in vitro detection of the antibody gene [sic] complexes possibly formed.

13. Method according to Claim 12, characterized in that the tissue or biological fluid is brought into contact with a mixture of polypeptide molecules responding [sic] to one of Claims 1 to 6 and other molecules originating from antigens of the sporozoite stage, namely LSA-1, SALSA or STARP.

14. Method of in vitro diagnosis of malaria in an individual likely to be infected by P. falciparum,

characterized in that it comprises the bringing of a tissue or biological fluid taken from an individual into contact with antibodies according to Claim 11, under conditions permitting an immunological reaction in vitro between the said antibodies and the proteins specific to P. falciparum which are possibly present in the biological tissue, and the in vitro detection of the antigen/antibody complexes possibly formed.

15. Kit for the in vitro diagnosis of malaria according to Claim 12 or 13, characterized in that it comprises at least one or several molecules according to one of Claims 1 to 6,
the reagents for making up the appropriate medium for the reaction,
the reagents enabling the antigen/antibody complexes produced by the immunological reaction to be detected, it also being possible for these reagents to carry a label or to be capable of being recognized in their turn by a labelled reagent, more especially in the case where the abovementioned polypeptide molecule is not labelled.

16. Kit for the in vitro diagnosis of malaria, characrerized in that it comprises:
- antibodies according to Claim 11,
- the reagents for making up the appropriate medium for carrying out the immunological reaction,
- the reagents enabling the antigen/antibody complexes produced by the immunological reaction to be detectd, it also being possible for these reagents to carry a label or to be capable of being recognized in their turn by a labelled reagent, more especially in the case where the abovementioned antibodies are not labelled.

17. Use of a polypeptide molecule according to one of Claims 1 to 6 in the preparation of an antimalarial vaccine.

18. Use of one or more polyclonal or monoclonal antibodies according to Claim 11 for the preparation of

a medicinal product intended for the treatment of malaria.

19. Pharmaceutical composition containing as active substance one or more polyclonal or monoclonal antibodies according to Claim 11, in combination with an acceptable pharmaceutical vehicle.

20. Nucleic acid sequence, characterized by one of the following sequences:

(a) the linked succession of nucleotides as depicted in SEQ ID No. 1 of Figure 1, or

(b) the linked succession of nucleotides depicted in SEQ ID No. 2 of Figure 2,

(c) a linked succession displaying at least 70% homology with that of Figure 1 or of Figure 2, or

(d) a linked succession of nucleotides which are complementary to those presented in (a), (b) or (c).

21. Nucleic acid according to Claim 20, containing a sequence coding for a polypeptide molecule according to one of Claims 1 to 6.

22. Recombinant vector for the cloning of a nucleotide sequence according to Claim 20 or Claim 21 and/or the expression of a polypeptide encoded by the abovementioned sequence, containing the said sequence in one of the sites which is not essential for its replication, the said vector being, in particular, of the plasmid, cosmid or phage type.

23. Vector according to Claim 22, characterized in that it is a plasmid deposited at the CNCM under the No. I-1573 and referenced pK1.2.

**ABSTRACT**

Polypeptide molecules containing at least 10 consecutive amino acids of the amino acid sequence shown in figure 2, representing the LSA3 antigen, the following peptides being excluded:

- RDELFNELLNSVDVNGEVKENILEESQVNDDIFNSLVKSVQQEQQHNVEE
- VEESVEENDEESVEENVEENVENNDDGSVASSVEESIASSVDESIDSSIE-ENVAPTVEEIVAPTVEEIVAPSVVEKCAPSVEESVAPSVEESVAEMLKER (729S)
- RDELFNELLNSVDVNGEVKENILEESQVNDDIFNSLVKSVQQEQQHN
- DELFNELLNSVDVNGEVKENILEESQ, (NRI)
- LEESQVNDDIFSNSLVKSVQQEQQHNV, (NRII)
- VESVAPSVEESVAPSVEESVAENVESSV. (729RE)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6152 base pairs

(B) TYPE: nucleotide

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATTTATTTAT TTTTATTGTT TTATTTCTTT TTTTTCTTTA AATTGTATAT TTATAAATAT 60

TTTAAAAAGT TAGAAAATGA CAAATAGTAA TTACAAATCA AATAATAAAA CATATAATGA 120

AAATAATAAT GAAGAAATAA CTACCATATT TAATAGAACA AATATGAATC CGATAAAAAA 180

ATGTCATATC AGAGAAAAAA TAAATAAGTA CTTTTTTTTG ATCAAAATTT TGACATGCAC 240

CATTTTAATA TGGGCTGTAC AATATGATAA TAACGTAAGA TAAAAAACTA AATAATAAAT 300

ATAAATAAAA AAAAAAAAAA AAAAAAAAAA ATCAACTATA TAGTATGTAT AATATATATA 360

TATATATATA TATATATATA TATATATATA TATTTATTTT TATTTATTTA TTAATTTTTT 420

TTTTTTTATA TTATCTTTTT AGTCTGATAT AAACAAGAGT TGGAAAAAAA ATACGTATGT 480

AGATAAGAAA TTGAATAAAC TATTTAACAC AAGTTTAGGA GAATCTCAAG TAAATGGTGA 540

ATTAGCTAGT GAAGAAGTAA AGGAAAAAAT TCTTGACTTA TTAGAAGAAG GAAATAGATT 600

AACTGAAAGT GTAGATGATA ATAAAAATTT AGAAGAAGCC GAAGATATAA AGGAAAATAT 660

CTTATTAAGT AATATAGAAC AACCAAAAGA AAATATTATT GACAATTTAT TAAATAATAT 720

TGGACAAAAT TCAGAAAAAC AAGAAAGTGT ATCAGAAAAT GTACAAGTCA GTGATGAACT 780

TTTTAATGAA TTATTAAATA GTGTAGATGT TAATGGAGAA GTAAAAGAAA ATATTTTGGA 840

GGAAAGTCAA GTTAATGACC ATATTTTTAA TAGTTTAGTA AAAAGTGTTC AACAAGAACA 900

ACAACACAAT GTTGAAGAAA AAGTTGAAGA AAGTGTAGAA GAAAATGACG AAGAAAGTGT 960

AGAAGAAAAT GTAGAAGAAA ATGTAGAAGA AAATGACGAC GGAAGTGTAG CCTCAAGTGT 1020

TGAAGAAAGT ATAGCTTCAA GTGTTGATGA AAGTATAGAT TCAAGTATTG AAGAAAATGT 1080

AGCTCCAACT GTTGAAGAAA TCGTAGCTCC AAGTGTTGTA GAAAGTGTGG GTCCAAGTGT 1140

FIGURE 1/A

TGAAGAAAGT GTAGAAGAAA ATGTTGAAGA AAGTGTAGCT GAAAATGTTG AAGAAAGTGT 1200

AGCTGAAAAT GTTGAAGAAA GTGTAGCTGA AAATGTTGAA GAAAGTGTAG CTGAAAATGT 1260

TGAAGAAATC GTAGCTCCAA CTGTTGAAGA AATCGTAGCT CCAACTGTTG AAGAAATTGT 1320

AGCTCCAAGT GTTGTAGAAA GTGTGGCTCC AAGTGTTGAA GAAAGTGTAG AAGAAAATGT 1380

TGAAGAAAGT GTAGCTGAAA ATGTTGAAGA AAGTGTAGCT GAAAATGTTG AAGAAAGTGT 1440

AGCTGAAAAT GTTGAAGAAA GTGTAGCTGA AAATGTTGAA GAAAGTGTAG CTGAAAATGT 1500

TGAAGAAATC GTAGCTCCAA CTGTTGAAGA AATCGTAGCT CCAACTGTTG AAGAAATTGT 1560

AGCTCCAAGT GTTGTAGAAA GTGTGGCTCC AAGTGTTGAA GAAAGTGTAG AAGAAAATGT 1620

TGAAGAAAGT GTAGCTGAAA ATGTTGAAGA AAGTGTAGCT GAAAATGTTG AAGAAAGTGT 1680

AGCTGAAAAT GTTGAAGAAA GTGTAGCTGA AAATGTTGAA GAAAGTGTAG CTGAAAATGT 1740

TGAAGAAAGT GTAGCTGAAA ATGTTGAAGA AAGTGTAGCT GAAAATGTTG AAGAAATCGT 1800

AGCTCCAACT GTTGAAGAAA TCGTAGCTCC AACTGTTGAA GAAATTGTAG CTCCAAGTGT 1860

TCTAGAAAGT GTGGCTCCAA GTGTTGAAGA AAGTGTAGAA GAAAATGTTG AAGAAAGTGT 1920

AGCTGAAAAT GTTGAAGAAA GTGTAGCTGA AAATGTTGAA GAAAGTGTAG CTGAAAATGT 1980

TGAAGAAAGT GTAGCTGAAA ATGTTGAAGA AATCGTAGCT CCAACTGTTG AAGAAATCGT 2040

AGCTCCAACT GTTGAAGAAA TTGTAGCTCC AAGTGTTGTA GAAAGTGTGG CTCCAAGTGT 2100

TGAAGAAAGT GTAGAAGAAA ATGTTGAAGA AAGTGTAGCT GAAAATGTTG AAGAAAGTGT 2160

AGCTGAAAAT GTTGAAGAAA GTGTAGCTGA AAATGTTGAA GAAATCGTAG CTCCAACTGT 2220

TGAAGAAATC GTAGCTCCAA CTGTTGAAGA AATTGTAGCT CCAAGTGTTG TAGAAAGTGT 2280

CGCTCCAAGT GTTGAAGAAA GTGTAGAAGA AAATGTTGAA GAAAGTGTAG CTGAAAATGT 2340

TGAAGAAAGT GTAGCTGAAA ATGTTGAAGA AAGTGTAGCT GAAAATGTTG AAGAAAGTGT 2400

AGCTGAAAAT GTTGAAGAAA TCGTAGCTCC AACTGTTGAA GAAATCGTAG CTCCAACTGT 2460

TGAAGAAATT GTAGCTCCAA GTGTTGTAGA AAGTGTGGCT CCAACTGTTG AAGAAAGTGT 2520

AGAAGAAAAT GTTGAAGAAA GTGTAGCTGA AAATGTTGAA GAAACTGTAG CTGAAAATGT 2580

TGAAGAAAGT GTAGCTGAAA ATGTTGAAGA AACTGTAGCT CCAACTGTTG AAGAAATTGT 2640

AGCTCCAAGT GTTGAAGAAA GTGTAGCTCC AAGTGTTGAA GAAAGTGTTG CTGAAAACGT 2700

TGCAACAAAT TTATCAGACA ATGTTTTAAG TAATTTATTA CGTGGTATCG AAACTGAGGA 2760

FIGURE 1/B

AATAAAGGAC AGTATATTAA ATGAGATAGA AGAAGTAAAA GAAAATGTAG TCACGACAAT 2820

ACTAGAAAAC GTAGAAGAAA CTACAGCTGA AAGTGTAACT ACTTTTAGTA ACATATTAGA 2880

GGAGATACAA GAAAATACTA TTACTAATGA TACTATAGAG GAAAAATTAG AAGAACTCCA 2940

CGAAAATGTA TTAAGTGCCC CTTTAGAAAA TACCCAAAGT GAAGAGGAAA AGAAAGAAGT 3000

AATAGATGTA ATTGAAGAAG TAAAAGAAGA GGTCGCTACC ACTTTAATAG AAACTGTGGA 3060

AGAGGCAGAA GAAAAGACGG CAAATACAAT TACGGAAATA TTTGAAAATT TAGAAGAAAA 3120

TGCAGTAGAA AGTAATGAAA ATGTTGCAGA CAATTTAGAG AAATTAAACG AAACTGTATT 3180

TAATACTGTA TTAGATAAAG TAGAGGAAAC AGTAGAAATT AGCGGAGAAA GTTTAGAAAA 3240

CAATGAAATG GATAAAGCAT TTTTTAGTGA AATATTTGAT AATGTAAAAG GAATACAAGA 3300

AAATTTATTA ACAGGTATGT TTCGAAGTAT AGAAACCAGT ATAGTAATCC AATCAGAAGA 3360

AAAGGTTGAT TTGAATGAAA ATGTGGTTAG TTCGATTTTA GATAATATAG AAAATATGAA 3420

AGAAGGTTTA TTAAATAAAT TAGAAAATAT TTCAAGTACT GAAGCTGTTC AAGAAACTGT 3480

AACTCAACAT GTAGAACAAA ATGTATATGT CGATGTTGAT GTTCCTGCTA TGAAAGATCA 3540

ATTTTTAGGA ATATTAAATG ACGCAGGAGG GTTGAAAGAA ATGTTTTTTA ATTTGGAAGA 3600

TGTATTTAAA AGTGAAAGTG ATGTAATTAC TGTAGAAGAA ATTAACGATG AACCGGTTCA 3660

AAAAGAGCTA GAAAAAGAAA CTGTTAGTAT TATTGAAGAA ATGGAAGAAA ATATTGTAGA 3720

TGTATTAGAG GAAGAAAAAG AAGATTTAAC AGACAAGATC ATAGATGCAG TAGAAGAATC 3780

CATAGAAATA TGTTCAGATT CTAAAGAAGA AACTGAATCT ATTAAAGATA AAGAAAAAGA 3840

TGTTTCACTA GTTGTTGAAG AAGTTCAAGA CAATGATATC GATGAAAGTG TTGAGAAAGT 3900

TTTAGAATTG AAAAATATGG AAGAGGAGTT AATGAACGAT GCTGTTGAAA TAAATGACAT 3960

TACTAGCAAA CTTATTGAAG AAACTCAAGA GTTAAATGAA GTAGAAGCAG ATTTAATAAA 4020

AGATATGGAA AAATTAAAAG AATTAGAAAA AGCATTATCA GAAGATTCTA AAGAAATAAT 4080

AGATGCAAAA GATGATACAT TAGAAAAAGT TATTCAACAG GAACATGATA TAACGACGAC 4140

GTTGGATGAA GTTGTAGAAT TAAAAGATGT CGAAGAAGAC AAGATCGAAA AAGTATCTGA 4200

TTTAAAAGAT CTTGAAGAAG ATATATTAAA AGAAGTAAAA GAAATCAAAG AAGTTGAAAG 4260

TGAAATTTTA GAAGATTATA AAGAATTAAA AACTATTGAA ACAGATATTT TAGAAGAGAA 4320

AAAAGAAATA GAAAAAGATC ATTTTGAAAA ATTCGAAGAA GAAGCTGAAG AAATAAAAGA 4380

FIGURE 1/C

TCTTGAAGCA GATATATTAA AAGAAGTATC TTCATTAGAA GTTGAAGAAG AAAAAAAATT 4440

AGAAGAAGTA CACGAATTAA AAGAAGAGGT AGAACATATA ATAAGTGGTG ATGCGCATAT 4500

AAAAGGTTTG GAAGAAGATG ATTTAGAAGA AGTAGATGAT TTAAAAGGAA GTATATTAGA 4560

CATGTTAAAG GGAGATATGG AATTAGCGGA TATCGATAAG GAAAGTTTAG AAGATGTAAC 4620

AACAAAACTT CGAGAAAGAG TTGAATCCTT AAAAGATGTT TTATCTAGTG CATTAGGCAT 4680

GGATGAAGAA CAAATGAAAA CAAGAAAAAA ACCTCAAAGA CGTAAGTTGC AAGAAGTATT 4740

ATTAAAAGAA GACGTTAAAG AAGAACCAAA GAAAAAAATA AGAAAAAAGA AAGTAAGGTT 4800

TGATATTAAG GATAAGGAAC CAAAAGATGA AATAGTAGAA GTTGAAATGA AAGATGAAGA 4860

TATAGAAGAA GATGTAGAAG AAGATATAGA AGAAGATATA GAAGAAGATA AAGTTGAAGA 4920

TATAGATGAA GATATAGATG AAGATATAGC TGAAGACAAA GATGAAGTTA TAGATTTAAT 4980

AGTCCAAAAA GAGAAACCCA TTGAAAAGGT TAAAGCGAAA AAGAAAAAAT TAGAAAAAAA 5040

AGTTGAAGAA GGTGTTAGTG GTCTTAAAAA ACACGTAGAC GAAGTAATGA AATATGTTCA 5100

AAAAATTCAT AAAGAAGTTG ATAAAGAAGT ATCTAAAGCT TTAGAATCAA AAAATGATGT 5160

TACTAATGTT TTAAAAACAA ATCAAGATTT TTTTAGTAAA GTTAAAAACT TCGTAAAAAA 5220

ATATAAAGTA TTTGCTCCAC CATTCATATC TGCCGTTGCA GCATTTGCAT GATATGTAGT 5280

TGGGTTCTTT ACATTTTCTT TATTTTCATC ATGTGTAACA ATAGCTTGTT CAAGTTAGTT 5340

ATTATCAAAA GTTGACAAAA CTATAAATAA AAATAAGGAG AGACCGTTTT ATTCATTTGT 5400

ATTTGATATC TTTAAGAATT TAAAACATTA TTTACAACAA ATGAAAGAAA AATTTAGTAA 5460

AGAAAAAAAT AATAATGTAA TAGAAGTAAC AAACAAAGCT GAGAAAAAAG GTAATGTACA 5520

GGTAACAAAT AAAACCGAGA AAACAACTAA AGTTGATAAA AATAATAAAG TACCGAAAAA 5580

AAGAAGAACG CAAAAATCAA AATAAAAAAT TGCAGAAGAG TGAAATGATT GGAGCGAACA 5640

ATAAAATTAA TCGATAAAAA ATATAAAAAT GTATATATTA TGTAAATATA TATAAATAAA 5700

TAAATAAATA CATACATATA TATATATATA TATATGTATC TTTTTACAAA ATTTTAAAAT 5760

TTTAAAATTT ATATATATTA ATATTTATAT TTTTCCATAT ATAATTTTAT TTTCAATATT 5820

TTATTTTTAA TTATAAATGT TTTTTACAGA GTTTATGTTT TTTAATTAAT ATATAGATTT 5880

CTGTAAGAAA CTGTATATTA TTCATACGAT ATATGTAATA TTAATTATTT GTGTTTTATT 5940

AAAATTTATA TTATATAATA TATATATATA TATATATGTA TATATATTAG AAGATAAAAA 6000

FIGURE 1/D

TTTAGCTTAT TTTGCTTGTT ATGCAAATAA GCTTTTTTTT TTTTTTTTTT TTTTTTTTTC 6060

ATATAAACGA TGTTTAATTT TTAATTTTTA ATATTTTATA TAAAATATTT TTCCTAAAAA 6120

AAAAAAAAAT TAAAAAAAAC TTATATTTCG AA 6152

FIGURE 1/E

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5361 base pairs

(B) TYPE: nucleotide

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..5361

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
ATG ACA AAT AGT AAT TAC AAA TCA AAT AAT AAA ACA TAT AAT GAA AAT      48
Met Thr Asn Ser Asn Tyr Lys Ser Asn Asn Lys Thr Tyr Asn Glu Asn
 1               5                  10                  15

AAT AAT GAA CAA ATA ACT ACC ATA TTT AAT AGA ACA AAT ATG AAT CCC      96
Asn Asn Glu Gln Ile Thr Thr Ile Phe Asn Arg Thr Asn Met Asn Pro
            20                  25                  30

ATA AAA AAA TGT CAT ATG AGA GAA AAA ATA AAT AAG TAC TTT TTT TTC     144
Ile Lys Lys Cys His Met Arg Glu Lys Ile Asn Lys Tyr Phe Phe Leu
        35                  40                  45

ATC AAA ATT TTG ACA TGC ACC ATT TTA ATA TGG GCT GTA CAA TAT GAT     192
Ile Lys Ile Leu Thr Cys Thr Ile Leu Ile Trp Ala Val Gln Tyr Asp
    50                  55                  60

AAT AAC TCT GAT ATA AAC AAG AGT TGG AAA AAA AAT ACG TAT GTA GAT     240
Asn Asn Ser Asp Ile Asn Lys Ser Trp Lys Lys Asn Thr Tyr Val Asp
65                  70                  75                  80

AAG AAA TTG AAT AAA CTA TTT AAC AGA AGT TTA GGA GAA TCT CAA GTA     288
Lys Lys Leu Asn Lys Leu Phe Asn Arg Ser Leu Gly Glu Ser Gln Val
                85                  90                  95

AAT GGT GAA TTA GCT AGT GAA GAA GTA AAG GAA AAA ATT CTT GAC TTA     336
Asn Gly Glu Leu Ala Ser Glu Glu Val Lys Glu Lys Ile Leu Asp Leu
            100                 105                 110

TTA GAA GAA GGA AAT AGA TTA ACT GAA AGT GTA GAT GAT AAT AAA AAT     384
Leu Glu Glu Gly Asn Thr Leu Thr Glu Ser Val Asp Asp Asn Lys Asn
        115                 120                 125
```

FIGURE 2/A

```
TTA GAA GAA GCC GAA GAT ATA AAG GAA AAT ATC TTA TTA AGT AAT ATA          432
Leu Glu Glu Ala Glu Asp Ile Lys Glu Asn Ile Leu Leu Ser Asn Ile
    130                 135                 140

GAA GAA CCA AAA GAA AAT ATT ATT GAC AAT TTA TTA AAT AAT ATT GGA          480
Glu Glu Pro Lys Glu Asn Ile Ile Asp Asn Leu Leu Asn Asn Ile Gly
145                 150                 155                 160

CAA AAT TCA GAA AAA CAA GAA AGT GTA TCA GAA AAT GTA CAA GTC AGT          528
Gln Asn Ser Glu Lys Gln Glu Ser Val Ser Glu Asn Val Gln Val Ser
                165                 170                 175

GAT GAA CTT TTT AAT GAA TTA TTA AAT AGT GTA GAT GTT AAT GGA GAA          576
Asp Glu Leu Phe Asn Glu Leu Leu Asn Ser Val Asp Val Asn Gly Glu
            180                 185                 190

GTA AAA GAA AAT ATT TTG GAG GAA AGT CAA GTT AAT GAC GAT ATT TTT          624
Val Lys Glu Asn Ile Leu Glu Glu Ser Gln Val Asn Asp Asp Ile Phe
        195                 200                 205

AAT AGT TTA GTA AAA AGT GTT CAA CAA GAA CAA CAA CAC AAT GTT GAA          672
Asn Ser Leu Val Lys Ser Val Gln Gln Glu Gln Gln His Asn Val Glu
    210                 215                 220

GAA AAA GTT GAA GAA AGT GTA GAA GAA AAT GAC GAA GAA AGT GTA GAA          720
Glu Lys Val Glu Glu Ser Val Glu Glu Asn Asp Glu Glu Ser Val Glu
225                 230                 235                 240

GAA AAT GTA GAA GAA AAT GTA GAA GAA AAT GAC GAC GGA AGT GTA GCC          768
Glu Asn Val Glu Glu Asn Val Glu Glu Asn Asp Asp Gly Ser Val Ala
                245                 250                 255

TCA AGT GTT GAA GAA AGT ATA GCT TCA AGT GTT GAT GAA AGT ATA GAT          816
Ser Ser Val Glu Glu Ser Ile Ala Ser Ser Val Asp Glu Ser Ile Asp
            260                 265                 270

TCA AGT ATT GAA GAA AAT GTA CCT CCA ACT GTT GAA GAA ATC GTA GCT          864
Ser Ser Ile Glu Glu Asn Val Ala Pro Thr Val Glu Glu Ile Val Ala
        275                 280                 285

CCA AGT GTT GTA GAA AGT GTC GCT CCA AGT GTT GAA GAA AGT GTA GAA          912
Pro Ser Val Val Glu Ser Val Ala Pro Ser Val Glu Glu Ser Val Glu
    290                 295                 300

GAA AAT GTT GAA GAA AGT GTA GCT GAA AAT GTT GAA GAA AGT GTA GCT          960
Glu Asn Val Glu Glu Ser Val Ala Glu Asn Val Glu Glu Ser Val Ala
305                 310                 315                 320

GAA AAT GTT CAA GAA AGT GTA CCT GAA AAT GTT GAA GAA AGT GTA GCT         1008
Glu Asn Val Glu Glu Ser Val Ala Glu Asn Val Glu Glu Ser Val Ala
                325                 330                 335
```

FIGURE 2/B

```
GAA AAT GTT GAA GAA ATC GTA GCT CCA ACT GTT GAA GAA ATC GTA GCT      1056
Glu Asn Val Glu Glu Ile Val Ala Pro Thr Val Glu Glu Ile Val Ala
            340                 345                 350

CCA ACT GTT GAA GAA ATT GTA GCT CCA AGT GTT GTA GAA AGT GTG GCT      1104
Pro Thr Val Glu Glu Ile Val Ala Pro Ser Val Val Glu Ser Val Ala
        355                 360                 365

CCA AGT GTT GAA GAA AGT GTA GAA GAA AAT GTT GAA GAA AGT GTA GCT      1152
Pro Ser Val Glu Glu Ser Val Glu Glu Asn Val Glu Glu Ser Val Ala
    370                 375                 380

GAA AAT GTT GAA GAA AGT GTA GCT GAA AAT GTT GAA GAA AGT GTA GCT      1200
Glu Asn Val Glu Glu Ser Val Ala Glu Asn Val Glu Glu Ser Val Ala
385                 390                 395                 400

GAA AAT GTT GAA GAA AGT GTA GCT GAA AAT GTT GAA GAA AGT GTA GCT      1248
Glu Asn Val Glu Glu Ser Val Ala Glu Asn Val Glu Glu Ser Val Ala
                405                 410                 415

GAA AAT GTT GAA GAA ATC GTA GCT CCA ACT GTT GAA GAA ATC GTA GCT      1296
Glu Asn Val Glu Glu Ile Val Ala Pro Thr Val Glu Glu Ile Val Ala
            420                 425                 430

CCA ACT GTT GAA GAA ATT GTA GCT CCA AGT GTT GTA GAA AGT GTG GCT      1344
Pro Thr Val Glu Glu Ile Val Ala Pro Ser Val Val Glu Ser Val Ala
        435                 440                 445

CCA AGT GTT GAA GAA AGT GTA GAA GAA AAT GTT GAA GAA AGT GTA GCT      1392
Pro Ser Val Glu Glu Ser Val Glu Glu Asn Val Glu Glu Ser Val Ala
    450                 455                 460

GAA AAT GTT GAA GAA AGT GTA GCT GAA AAT GTT GAA GAA AGT GTA GCT      1440
Glu Asn Val Glu Glu Ser Val Ala Glu Asn Val Glu Glu Ser Val Ala
465                 470                 475                 480

GAA AAT GTT GAA GAA AGT GTA GCT GAA AAT GTT GAA GAA AGT GTA GCT      1488
Glu Asn Val Glu Glu Ser Val Ala Glu Asn Val Glu Glu Ser Val Ala
                485                 490                 495

GAA AAT GTT GAA GAA AGT GTA GCT GAA AAT GTT GAA GAA AGT GTA GCT      1536
Glu Asn Val Glu Glu Ser Val Ala Glu Asn Val Glu Glu Ser Val Ala
            500                 505                 510

GAA AAT GTT GAA GAA ATC GTA GCT CCA ACT GTT GAA GAA ATC GTA GCT      1584
Glu Asn Val Glu Glu Ile Val Ala Pro Thr Val Glu Glu Ile Val Ala
        515                 520                 525

CCA ACT GTT GAA GAA ATT GTA GCT CCA AGT GTT GTA GAA AGT GTG GCT      1632
Pro Thr Val Glu Glu Ile Val Ala Pro Ser Val Val Glu Ser Val Ala
    530                 535                 540
```

FIGURE 2/C

```
CCA AGT GTT GAA GAA AGT GTA GAA GAA AAT GTT GAA GAA AGT GTA GCT      1680
Pro Ser Val Glu Glu Ser Val Glu Glu Asn Val Glu Glu Ser Val Ala
545                 550                 555                 560

GAA AAT GTT GAA GAA AGT GTA GCT GAA AAT GTT GAA GAA AGT GTA GCT      1728
Glu Asn Val Glu Glu Ser Val Ala Glu Asn Val Glu Glu Ser Val Ala
                565                 570                 575

GAA AAT GTT GAA GAA AGT GTA GCT GAA AAT GTT GAA GAA ATC GTA GCT      1776
Glu Asn Val Glu Glu Ser Val Ala Glu Asn Val Glu Glu Ile Val Ala
            580                 585                 590

CCA ACT GTT GAA GAA ATC GTA GCT CCA ACT GTT GAA GAA ATT GTA GCT      1824
Pro Thr Val Glu Glu Ile Val Ala Pro Thr Val Glu Glu Ile Val Ala
        595                 600                 605

CCA AGT GTT GTA GAA AGT GTG GCT CCA AGT GTT GAA GAA AGT GTA GAA      1872
Pro Ser Val Val Glu Ser Val Ala Pro Ser Val Glu Glu Ser Val Glu
    610                 615                 620

GAA AAT GTT GAA GAA AGT GTA GCT GAA AAT GTT GAA GAA AGT GTA GCT      1920
Glu Asn Val Glu Glu Ser Val Ala Glu Asn Val Glu Glu Ser Val Ala
625                 630                 635                 640

GAA AAT GTT GAA GAA AGT GTA GCT GAA AAT GTT GAA GAA ATC GTA GCT      1968
Glu Asn Val Glu Glu Ser Val Ala Glu Asn Val Glu Glu Ile Val Ala
                645                 650                 655

CCA ACT GTT GAA GAA ATC GTA GCT CCA ACT GTT GAA GAA ATT GTA GCT      2016
Pro Thr Val Glu Glu Ile Val Ala Pro Thr Val Glu Glu Ile Val Ala
            660                 665                 670

CCA AGT GTT GTA GAA AGT GTG GCT CCA AGT GTT GAA GAA AGT GTA GAA      2064
Pro Ser Val Val Glu Ser Val Ala Pro Ser Val Glu Glu Ser Val Glu
        675                 680                 685

GAA AAT GTT GAA GAA AGT GTA GCT GAA AAT GTT GAA GAA AGT GTA GCT      2112
Glu Asn Val Glu Glu Ser Val Ala Glu Asn Val Glu Glu Ser Val Ala
    690                 695                 700

GAA AAT GTT GAA GAA AGT GTA GCT GAA AAT GTT GAA GAA AGT GTA GCT      2160
Glu Asn Val Glu Glu Ser Val Ala Glu Asn Val Glu Glu Ser Val Ala
705                 710                 715                 720

GAA AAT GTT GAA GAA ATC GTA GCT CCA ACT GTT GAA GAA ATC GTA GCT      2208
Glu Asn Val Glu Glu Ile Val Ala Pro Thr Val Glu Glu Ile Val Ala
                725                 730                 735

CCA ACT GTT GAA GAA ATT GTA GCT CCA AGT GTT GTA GAA AGT GTG GCT      2256
Pro Thr Val Glu Glu Ile Val Ala Pro Ser Val Val Glu Ser Val Ala
            740                 745                 750
```

FIGURE 2/D.

```
CCA AGT GTT GAA GAA AGT GTA GAA GAA AAT GTT GAA GAA AGT GTA GCT      2304
Pro Ser Val Glu Glu Ser Val Glu Glu Asn Val Glu Glu Ser Val Ala
        755                 760                 765

GAA AAT GTT GAA GAA AGT GTA GCT GAA AAT GTT GAA GAA AGT GTA GCT      2352
Glu Asn Val Glu Glu Ser Val Ala Glu Asn Val Glu Glu Ser Val Ala
    770                 775                 780

GAA AAT GTT GAA GAA AGT GTA GCT CCA ACT GTT GAA GAA ATT GTA GCT      2400
Glu Asn Val Glu Glu Ser Val Ala Pro Thr Val Glu Glu Ile Val Ala
785                 790                 795                 800

CCA AGT GTT GAA GAA ACT GTA GCT CCA AGT GTT GAA GAA AGT GTT GCT      2448
Pro Ser Val Glu Glu Ser Val Ala Pro Ser Val Glu Glu Ser Val Ala
                805                 810                 815

GAA AAC GTT GCA ACA AAT TTA TCA GAC AAT CTT TTA AGT AAT TTA TTA      2496
Glu Asn Val Ala Thr Asn Leu Ser Asp Asn Leu Leu Ser Asn Leu Leu
            820                 825                 830

GGT GGT ATC GAA ACT GAG GAA ATA AAG GAC AGT ATA TTA AAT GAG ATA      2544
Gly Gly Ile Glu Thr Glu Glu Ile Lys Asp Ser Ile Leu Asn Glu Ile
        835                 840                 845

GAA GAA GTA AAA GAA AAT GTA GTC ACC ACA ATA CTA GAA AAC GTA GAA      2592
Glu Glu Val Lys Glu Asn Val Val Thr Thr Ile Leu Glu Asn Val Glu
    850                 855                 860

GAA ACT ACA GCT GAA AGT GTA ACT ACT TTT AGT AAC ATA TTA GAG GAG      2640
Glu Thr Thr Ala Glu Ser Val Thr Thr Phe Ser Asn Ile Leu Glu Glu
865                 870                 875                 880

ATA CAA GAA AAT ACT ATT ACT AAT GAT ACT ATA GAG GAA AAA TTA GAA      2688
Ile Gln Glu Asn Thr Ile Thr Asn Asp Thr Ile Glu Glu Lys Leu Glu
                885                 890                 895

GAA CTC CAC GAA AAT GTA TTA AGT GCC GCT TTA GAA AAT ACC CAA AGT      2736
Glu Leu His Glu Asn Val Leu Ser Ala Ala Leu Glu Asn Thr Gln Ser
            900                 905                 910

GAA GAG GAA AAG AAA GAA GTA ATA GAT GTA ATT GAA GAA GTA AAA GAA      2784
Glu Glu Glu Lys Lys Glu Val Ile Asp Val Ile Glu Glu Val Lys Glu
        915                 920                 925

GAG GTC GCT ACC ACT TTA ATA GAA ACT GTG GAA CAG GCA GAA GAA AAG      2832
Glu Val Ala Thr Thr Leu Ile Glu Thr Val Glu Gln Ala Glu Glu Lys
    930                 935                 940

AGC GCA AAT ACA ATT ACG GAA ATA TTT GAA AAT TTA GAA GAA AAT GCA      2880
Ser Ala Asn Thr Ile Thr Glu Ile Phe Glu Asn Leu Glu Glu Asn Ala
945                 950                 955                 960
```

FIGURE 2/E

```
GTA GAA AGT AAT GAA AAT GTT GCA GAG AAT TTA GAG AAA TTA AAC GAA      2928
Val Glu Ser Asn Glu Asn Val Ala Glu Asn Leu Glu Lys Leu Asn Glu
                965                 970                 975

ACT GTA TTT AAT ACT GTA TTA GAT AAA GTA GAG GAA ACA GTA GAA ATT      2976
Thr Val Phe Asn Thr Val Leu Asp Lys Val Glu Glu Thr Val Glu Ile
            980                 985                 990

AGC GGA GAA AGT TTA GAA AAC AAT GAA ATG GAT AAA GCA TTT TTT AGT      3024
Ser Gly Glu Ser Leu Glu Asn Asn Glu Met Asp Lys Ala Phe Phe Ser
        995                 1000                1005

GAA ATA TTT GAT AAT GTA AAA GGA ATA CAA GAA AAT TTA TTA ACA GGT      3072
Glu Ile Phe Asp Asn Val Lys Gly Ile Gln Glu Asn Leu Leu Thr Gly
    1010                1015                1020

ATG TTT CGA AGT ATA GAA ACC AGT ATA GTA ATC CAA TCA GAA GAA AAG      3120
Met Phe Arg Ser Ile Glu Thr Ser Ile Val Ile Gln Ser Glu Glu Lys
1025                1030                1035                1040

GTT GAT TTG AAT GAA AAT GTG GTT AGT TCC ATT TTA GAT AAT ATA GAA      3168
Val Asp Leu Asn Glu Asn Val Val Ser Ser Ile Leu Asp Asn Ile Glu
                1045                1050                1055

AAT ATG AAA GAA GGT TTA TTA AAT AAA TTA GAA AAT ATT TCA AGT ACT      3216
Asn Met Lys Glu Gly Leu Leu Asn Lys Leu Glu Asn Ile Ser Ser Thr
            1060                1065                1070

GAA GGT GTT CAA GAA ACT GTA ACT GAA CAT GTA GAA CAA AAT GTA TAT      3264
Glu Gly Val Gln Glu Thr Val Thr Glu His Val Glu Gln Asn Val Tyr
        1075                1080                1085

GTG GAT GTT GAT GTT CCT GCT ATG AAA GAT CAA TTT TTA GGA ATA TTA      3312
Val Asp Val Asp Val Pro Ala Met Lys Asp Gln Phe Leu Gly Ile Leu
    1090                1095                1100

AAT GAG GCA GGA GGC TTG AAA GAA ATG TTT TTT AAT TTG GAA GAT GTA      3360
Asn Glu Ala Gly Gly Leu Lys Glu Met Phe Phe Asn Leu Glu Asp Val
1105                1110                1115                1120

TTT AAA AGT GAA AGT GAT GTA ATT ACT GTA GAA GAA ATT AAG GAT GAA      3408
Phe Lys Ser Glu Ser Asp Val Ile Thr Val Glu Glu Ile Lys Asp Glu
                1125                1130                1135

CCG GTT CAA AAA GAG GTA GAA AAA GAA ACT GTT AGT ATT ATT GAA GAA      3456
Pro Val Gln Lys Glu Val Glu Lys Glu Thr Val Ser Ile Ile Glu Glu
            1140                1145                1150

ATG GAA GAA AAT ATT GTA GAT GTA TTA GAG GAA GAA AAA GAA GAT TTA      3504
Met Glu Glu Asn Ile Val Asp Val Leu Glu Glu Glu Lys Glu Asp Leu
        1155                1160                1165
```

FIGURE 2/F

```
ACA GAC AAG ATG ATA GAT GCA GTA GAA GAA TCC ATA GAA ATA TCT TCA      3552
Thr Asp Lys Met Ile Asp Ala Val Glu Glu Ser Ile Glu Ile Ser Ser
    1170                1175                1180

GAT TCT AAA GAA GAA ACT GAA TCT ATT AAA GAT AAA GAA AAA GAT GTT      3600
Asp Ser Lys Glu Glu Thr Glu Ser Ile Lys Asp Lys Glu Lys Asp Val
1185                1190                1195                1200

TCA CTA GTT GTT GAA GAA GTT CAA GAC AAT GAT ATG GAT GAA AGT GTT      3648
Ser Leu Val Val Glu Glu Val Gln Asp Asn Asp Met Asp Glu Ser Val
                1205                1210                1215

GAG AAA GTT TTA GAA TTG AAA AAT ATG GAA GAG GAG TTA ATG AAG GAT      3696
Glu Lys Val Leu Glu Leu Lys Asn Met Glu Glu Glu Leu Met Lys Asp
            1220                1225                1230

GCT GTT GAA ATA AAT GAC ATT ACT AGC AAA CTT ATT GAA GAA ACT CAA      3744
Ala Val Glu Ile Asn Asp Ile Thr Ser Lys Leu Ile Glu Glu Thr Gln
        1235                1240                1245

GAG TTA AAT GAA GTA GAA GCA GAT TTA ATA AAA GAT ATG GAA AAA TTA      3792
Glu Leu Asn Glu Val Glu Ala Asp Leu Ile Lys Asp Met Glu Lys Leu
    1250                1255                1260

AAA GAA TTA GAA AAA GCA TTA TCA GAA GAT TCT AAA GAA ATA ATA GAT      3840
Lys Glu Leu Glu Lys Ala Leu Ser Glu Asp Ser Lys Glu Ile Ile Asp
1265                1270                1275                1280

GCA AAA GAT GAT ACA TTA GAA AAA GTT ATT GAA GAG GAA CAT GAT ATA      3888
Ala Lys Asp Asp Thr Leu Glu Lys Val Ile Glu Glu Glu His Asp Ile
                1285                1290                1295

ACG ACG ACG TTG GAT GAA GTT GTA GAA TTA AAA GAT GTC GAA GAA GAC      3936
Thr Thr Thr Leu Asp Glu Val Val Glu Leu Lys Asp Val Glu Glu Asp
            1300                1305                1310

AAG ATC GAA AAA GTA TCT GAT TTA AAA GAT CTT GAA GAA GAT ATA TTA      3984
Lys Ile Glu Lys Val Ser Asp Leu Lys Asp Leu Glu Glu Asp Ile Leu
        1315                1320                1325

AAA GAA GTA AAA GAA ATC AAA GAA CTT GAA AGT GAA ATT TTA GAA GAT      4032
Lys Glu Val Lys Glu Ile Lys Glu Leu Glu Ser Glu Ile Leu Glu Asp
    1330                1335                1340

TAT AAA GAA TTA AAA ACT ATT GAA ACA GAT ATT TTA GAA GAG AAA AAA      4080
Tyr Lys Glu Leu Lys Thr Ile Glu Thr Asp Ile Leu Glu Glu Lys Lys
1345                1350                1355                1360

GAA ATA GAA AAA GAT CAT TTT GAA AAA TTC GAA GAA GAA GCT GAA GAA      4128
Glu Ile Glu Lys Asp His Phe Glu Lys Phe Glu Glu Glu Ala Glu Glu
                1365                1370                1375
```

FIGURE 2/G

```
ATA AAA GAT CTT GAA GCA GAT ATA TTA AAA GAA GTA TCT TCA TTA GAA      4176
Ile Lys Asp Leu Glu Ala Asp Ile Leu Lys Glu Val Ser Ser Leu Glu
            1380                1385                1390

GTT GAA GAA GAA AAA AAA TTA GAA GAA GTA CAC GAA TTA AAA GAA GAG      4224
Val Glu Glu Glu Lys Lys Leu Glu Glu Val His Glu Leu Lys Glu Glu
        1395                1400                1405

GTA GAA CAT ATA ATA AGT GGT GAT GCG CAT ATA AAA GGT TTG GAA GAA      4272
Val Glu His Ile Ile Ser Gly Asp Ala His Ile Lys Gly Leu Glu Glu
    1410                1415                1420

GAT GAT TTA GAA GAA GTA GAT GAT TTA AAA GGA AGT ATA TTA GAC ATG      4320
Asp Asp Leu Glu Glu Val Asp Asp Leu Lys Gly Ser Ile Leu Asp Met
1425                1430                1435                1440

TTA AAG GGA GAT ATG GAA TTA GGC GAT ATG GAT AAG GAA AGT TTA GAA      4368
Leu Lys Gly Asp Met Glu Leu Gly Asp Met Asp Lys Glu Ser Leu Glu
                1445                1450                1455

GAT GTA AGA ACA AAA CTT GGA GAA AGA GTT GAA TCC TTA AAA GAT GTT      4416
Asp Val Thr Thr Lys Leu Gly Glu Arg Val Glu Ser Leu Lys Asp Val
            1460                1465                1470

TTA TCT AGT GCA TTA GGC ATG GAT GAA GAA CAA ATG AAA ACA AGA AAA      4464
Leu Ser Ser Ala Leu Gly Met Asp Glu Glu Gln Met Lys Thr Arg Lys
        1475                1480                1485

AAA GCT CAA AGA CCT AAG TTG GAA GAA GTA TTA TTA AAA GAA GAG GTT      4512
Lys Ala Gln Arg Pro Lys Leu Glu Glu Val Leu Leu Lys Glu Glu Val
    1490                1495                1500

AAA GAA GAA CCA AAG AAA AAA ATA ACA AAA AAG AAA GTA AGG TTT GAT      4560
Lys Glu Glu Pro Lys Lys Lys Ile Thr Lys Lys Lys Val Arg Phe Asp
1505                1510                1515                1520

ATT AAG GAT AAG GAA CCA AAA GAT GAA ATA GTA GAA GTT GAA ATG AAA      4608
Ile Lys Asp Lys Glu Pro Lys Asp Glu Ile Val Glu Val Glu Met Lys
                1525                1530                1535

GAT GAA GAT ATA GAA GAA GAT GTA GAA GAA GAT ATA GAA GAA GAT ATA      4656
Asp Glu Asp Ile Glu Glu Asp Val Glu Glu Asp Ile Glu Glu Asp Ile
            1540                1545                1550

GAA GAA GAT AAA GTT GAA GAT ATA GAT GAA GAT ATA GAT GAA GAT ATA      4704
Glu Glu Asp Lys Val Glu Asp Ile Asp Glu Asp Ile Asp Glu Asp Ile
        1555                1560                1565

GGT GAA GAC AAA GAT GAA GTT ATA GAT TTA ATA GTC CAA AAA GAG AAA      4752
Gly Glu Asp Lys Asp Glu Val Ile Asp Leu Ile Val Gln Lys Glu Lys
    1570                1575                1580
```

FIGURE 2/H

```
CGC ATT GAA AAG GTT AAA GCG AAA AAG AAA AAA TTA GAA AAA AAA GTT      4800
Arg Ile Glu Lys Val Lys Ala Lys Lys Lys Lys Leu Glu Lys Lys Val
1585                1590                1595                1600

GAA GAA GGT GTT AGT GGT CTT AAA AAA CAC GTA GAC GAA GTA ATG AAA      4848
Glu Glu Gly Val Ser Gly Leu Lys Lys His Val Asp Glu Val Met Lys
                1605                1610                1615

TAT GTT CAA AAA ATT GAT AAA GAA GTT GAT AAA GAA GTA TCT AAA GCT      4896
Tyr Val Gln Lys Ile Asp Lys Glu Val Asp Lys Glu Val Ser Lys Ala
            1620                1625                1630

TTA GAA TCA AAA AAT GAT GTT ACT AAT GTT TTA AAA CAA AAT CAA GAT      4944
Leu Glu Ser Lys Asn Asp Val Thr Asn Val Leu Lys Gln Asn Gln Asp
        1635                1640                1645

TTT TTT AGT AAA GTT AAA AAC TTC GTA AAA AAA TAT AAA GTA TTT GCT      4992
Phe Phe Ser Lys Val Lys Asn Phe Val Lys Lys Tyr Lys Val Phe Ala
    1650                1655                1660

GCA CGA TTC ATA TCT GCC GTT GCA GCA TTT GCA TCA TAT GTA GTT GGC      5040
Ala Pro Phe Ile Ser Ala Val Ala Ala Phe Ala Ser Tyr Val Val Gly
1665                1670                1675                1680

TTC TTT ACA TTT TCT TTA TTT TCA TCA TGT GTA ACA ATA GCT TCT TCA      5088
Phe Phe Thr Phe Ser Leu Phe Ser Ser Cys Val Thr Ile Ala Ser Ser
                1685                1690                1695

ACT TAC TTA TTA TCA AAA GTT GAC AAA ACT ATA AAT AAA AAT AAG GAG      5136
Thr Tyr Leu Leu Ser Lys Val Asp Lys Thr Ile Asn Lys Asn Lys Glu
            1700                1705                1710

AGA CCG TTT TAT TCA TTT GTA TTT GAT ATC TTT AAG AAT TTA AAA CAT      5184
Arg Pro Phe Tyr Ser Phe Val Phe Asp Ile Phe Lys Asn Leu Lys His
        1715                1720                1725

TAT TTA CAA CAA ATG AAA GAA AAA TTT AGT AAA GAA AAA AAT AAT AAT      5232
Tyr Leu Gln Gln Met Lys Glu Lys Phe Ser Lys Glu Lys Asn Asn Asn
    1730                1735                1740

GTA ATA GAA GTA ACA AAC AAA GCT CAG AAA AAA GGT AAT GTA CAG GTA      5280
Val Ile Glu Val Thr Asn Lys Ala Glu Lys Lys Gly Asn Val Gln Val
1745                1750                1755                1760

ACA AAT AAA ACC GAG AAA ACA ACT AAA GTT GAT AAA AAT AAT AAA GTA      5328
Thr Asn Lys Thr Glu Lys Thr Thr Lys Val Asp Lys Asn Asn Lys Val
                1765                1770                1775

CCG AAA AAA AGA AGA ACG CAA AAA TCA AAA TAA                          5361
Pro Lys Lys Arg Arg Thr Gln Lys Ser Lys *
            1780                1785
```

FIGURE 2/I

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1891 base pairs

(B) TYPE: nucleotide

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 2..1891

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```
T ACA TTA ACT GAA AGT GTA GAT GAT AAT AAA AAT TTA GAA GAA GCC                46
  Thr Leu Thr Glu Ser Val Asp Asp Asn Lys Asn Leu Glu Glu Ala
    1               5                  10                  15

GAA GAT ATA AAG GAA AAT ATC TTA TTA AGT AAT ATA GAA GAA GCA AAA              94
Glu Asp Ile Lys Glu Asn Ile Leu Leu Ser Asn Ile Glu Glu Pro Lys
            20                  25                  30

GAA AAT ATT ATT GAC AAT TTA TTA AAT AAT ATT GGA CAA AAT TCA GAA             142
Glu Asn Ile Ile Asp Asn Leu Leu Asn Asn Ile Gly Gln Asn Ser Glu
        35                  40                  45

AAA CAA GAA AGT GTA TCA GAA AAT GTA CAA GTC AGT GAT GAA CTT TTT             190
Lys Gln Glu Ser Val Ser Glu Asn Val Gln Val Ser Asp Glu Leu Phe
    50                  55                  60

AAT GAA TTA TTA AAT AGT GTA GAT GTT AAT GGA GAA GTA AAA GAA AAT             238
Asn Glu Leu Leu Asn Ser Val Asp Val Asn Gly Glu Val Lys Glu Asn
65                  70                  75

ATT TTG GAC GAA AGT CAA GTT AAT GAC GAT ATT TTT AAT AGT TTA GTA             286
Ile Leu Glu Glu Ser Gln Val Asn Asp Asp Ile Phe Asn Ser Leu Val
80                  85                  90                  95

AAA AGT GTT CAA CAA GAA CAA CAA CAC AAT GTT GAA GAA AAA GTT GAA             334
Lys Ser Val Gln Gln Glu Gln Gln His Asn Val Glu Glu Lys Val Glu
            100                 105                 110

GAA AGT GTA GAA GAA AAT GAC GAA GAA AGT GTA GAA GAA AAT GTA GAA             382
Glu Ser Val Glu Glu Asn Asp Glu Glu Ser Val Glu Glu Asn Val Glu
        115                 120                 125

GAA AAT GTA GAA GAA AAT GAC GAC GGA AGT GTA GCC TCA AGT GTT GAA             430
Glu Asn Val Glu Glu Asn Asp Asp Gly Ser Val Ala Ser Ser Val Glu
    130                 135                 140
```

FIGURE 3/A

```
GAA AGT ATA GCT TCA AGT GTT GAT GAA AGT ATA GAT TCA AGT ATT GAA        478
Glu Ser Ile Ala Ser Ser Val Asp Glu Ser Ile Asp Ser Ser Ile Glu
    145                 150                 155

GAA AAT GTA GCT CCA ACT GTT GAA GAA ATC GTA GCT CCA ACT GTT GAA        526
Glu Asn Val Ala Pro Thr Val Glu Glu Ile Val Ala Pro Thr Val Glu
160                 165                 170                 175

GAA ATT GTA GCT CCA AGT GTT GTA GAA AGT GTG GCT CCA AGT GTT GAA        574
Glu Ile Val Ala Pro Ser Val Val Glu Ser Val Ala Pro Ser Val Glu
            180                 185                 190

GAA AGT GTA GCT CCA AGT GTT GAA GAA AGT GTA GCT GAA AAT GTT GAA        622
Glu Ser Val Ala Pro Ser Val Glu Glu Ser Val Ala Glu Asn Val Glu
        195                 200                 205

GAA AGT GTA GCT GAA AAT GTT GAA GAA ATC GTA GCT CCA AGT GTT GAA        670
Glu Ser Val Ala Glu Asn Val Glu Glu Ile Val Ala Pro Ser Val Glu
    210                 215                 220

GAA AGT GTA GCT GAA AAT GTT GAA GAA AGT GTA GCT GAA AAT GTT GAA        718
Glu Ser Val Ala Glu Asn Val Glu Glu Ser Val Ala Glu Asn Val Glu
225                 230                 235

GAA AGT GTA GCT GAA AAT GTT GAA GAA AGT GTA GCT GAA AAT GTT GAA        766
Glu Ser Val Ala Glu Asn Val Glu Glu Ser Val Ala Glu Asn Val Glu
240                 245                 250                 255

GAA AGT GTA GCT GAA AAT GTT GAA GAA ATC GTA GCT CCA ACT GTT GAA        814
Glu Ser Val Ala Glu Asn Val Glu Glu Ile Val Ala Pro Thr Val Glu
            260                 265                 270

GAA AGT GTA GCT CCA ACT GTT GAA GAA ATT GTA GCT CCA ACT GTT GAA        862
Glu Ser Val Ala Pro Thr Val Glu Glu Ile Val Ala Pro Thr Val Glu
        275                 280                 285

GAA AGT GTA GCT CCA ACT GTT GAA GAA ATT GTA GTT CCA AGT GTT GAA        910
Glu Ser Val Ala Pro Thr Val Glu Glu Ile Val Val Pro Ser Val Glu
    290                 295                 300

GAA AGT GTA GCT CCA AGT GTT GAA GAA AGT GTA GCT GAA AAT GTT GAA        958
Glu Ser Val Ala Pro Ser Val Glu Glu Ser Val Ala Glu Asn Val Glu
305                 310                 315

GAA AGT GTA GCT GAA AAT GTT GAA GAA AGT GTA GCT GAA AAT GTT GAA       1006
Glu Ser Val Ala Glu Asn Val Glu Glu Ser Val Ala Glu Asn Val Glu
320                 325                 330                 335

GAA ACT GTA GCT GAA AAT GTT GAA GAA AGT GTA GCT GAA AAT GTT GAA       1054
Glu Ser Val Ala Glu Asn Val Glu Glu Ser Val Ala Glu Asn Val Glu
            340                 345                 350
```

FIGURE 3/B

```
GAA ATC GTA GCT CCA AGT GTT GAA GAA ATC GTA GCT CCA ACT GTT GAA      1102
Glu Ile Val Ala Pro Ser Val Glu Glu Ile Val Ala Pro Thr Val Glu
            355                 360                 365

GAA AGT GTT GCT GAA AAC GTT GCA ACA AAT TTA TCA GAC AAT GTT TTA      1150
Glu Ser Val Ala Glu Asn Val Ala Thr Asn Leu Ser Asp Asn Leu Leu
        370                 375                 380

AGT AAT TTA TTA GGT GGT ATC GAA ACT GAC GAA ATA AAG GAC AGT ATA      1198
Ser Asn Leu Leu Gly Gly Ile Glu Thr Glu Glu Ile Lys Asp Ser Ile
    385                 390                 395

TTA AAT GAG ATA GAA GAA GTA AAA GAA AAT GTA GTC ACC ACA ATA CTA      1246
Leu Asn Glu Ile Glu Glu Val Lys Glu Asn Val Val Thr Thr Ile Leu
400                 405                 410                 415

GAA AAA GTA GAA GAA ACT ACA GCT GAA AGT GTA ACT ACT TTT AGT AAT      1294
Glu Lys Val Glu Glu Thr Thr Ala Glu Ser Val Thr Thr Phe Ser Asn
                420                 425                 430

ATA TTA GAG GAG ATA CAA GAA AAT ACT ATT ACT AAT GAT ACT ATA GAG      1342
Ile Leu Glu Glu Ile Gln Glu Asn Thr Ile Thr Asn Asp Thr Ile Glu
            435                 440                 445

GAA AAA TTA GAA GAA CTC CAC GAA AAT GTA TTA AGT GCC GCT TTA GAA      1390
Glu Lys Leu Glu Glu Leu His Glu Asn Val Leu Ser Ala Ala Leu Glu
        450                 455                 460

AAT ACC CAA AGT GAA GAG GAA AAG AAA GAA GTA ATA GAT GTA ATT GAA      1438
Asn Thr Gln Ser Glu Glu Glu Lys Lys Glu Val Ile Asp Val Ile Glu
    465                 470                 475

GAA GTA AAA GAA GAG GTC GCT ACC ACT TTA ATA GAA ACT GTG GAA CAG      1486
Glu Val Lys Glu Glu Val Ala Thr Thr Leu Ile Glu Thr Val Glu Gln
480                 485                 490                 495

GCA GAA GAA GAG AGC GAA AGT ACA ATT ACG GAA ATA TTT GAA AAT TTA      1534
Ala Glu Glu Glu Ser Glu Ser Thr Ile Thr Glu Ile Phe Glu Asn Leu
                500                 505                 510

GAA GAA AAT GCA GTA GAA AGT AAT GAA AAA GTT GCA GAG AAT TTA GAG      1582
Glu Glu Asn Ala Val Glu Ser Asn Glu Lys Val Ala Glu Asn Leu Glu
            515                 520                 525

AAA TTA AAC GAA ACT GTA TTT AAT ACT GTA TTA GAT AAA GTA GAG GAA      1630
Lys Leu Asn Glu Thr Val Phe Asn Thr Val Leu Asp Lys Val Glu Glu
        530                 535                 540

ACA GTA GAA ATT AGC GGA GAA AGT TTA GAA AAC AAT GAA ATG GAT AAA      1678
Thr Val Glu Ile Ser Gly Glu Ser Leu Glu Asn Asn Glu Met Asp Lys
    545                 550                 555
```

FIGURE 3/C

```
GCA TTT TTT AGT GAA ATA TTT GAT AAT GTA AAA GGA ATA CAA GAA AAT          1726
Ala Phe Phe Ser Glu Ile Phe Asp Asn Val Lys Gly Ile Gln Glu Asn
560                 565                 570                 575

TTA TTA ACA GGT ATG TTT CGA AGT ATA GAA ACC AGT ATA GTA ATC CAA          1774
Leu Leu Thr Gly Met Phe Arg Ser Ile Glu Thr Ser Ile Val Ile Gln
                580                 585                 590

TCA GAA GAA AAG GTT GAT TTG AAT GAA AAT GTG GTT AGT TCG ATT TTA          1822
Ser Glu Glu Lys Val Asp Leu Asn Glu Asn Val Val Ser Ser Ile Leu
            595                 600                 605

GAT AAT ATA GAA AAT ATG AAA GAA GGT TTA TTA AAT AAA TTA GAA AAT          1870
Asp Asn Ile Glu Asn Met Lys Glu Gly Leu Leu Asn Lys Leu Glu Asn
        610                 615                 620

ATT TCA AGT ACT GAA GGC GAA                                             1891
Ile Ser Ser Thr Glu Gly Glu
    625                 630
```

FIGURE 3/D

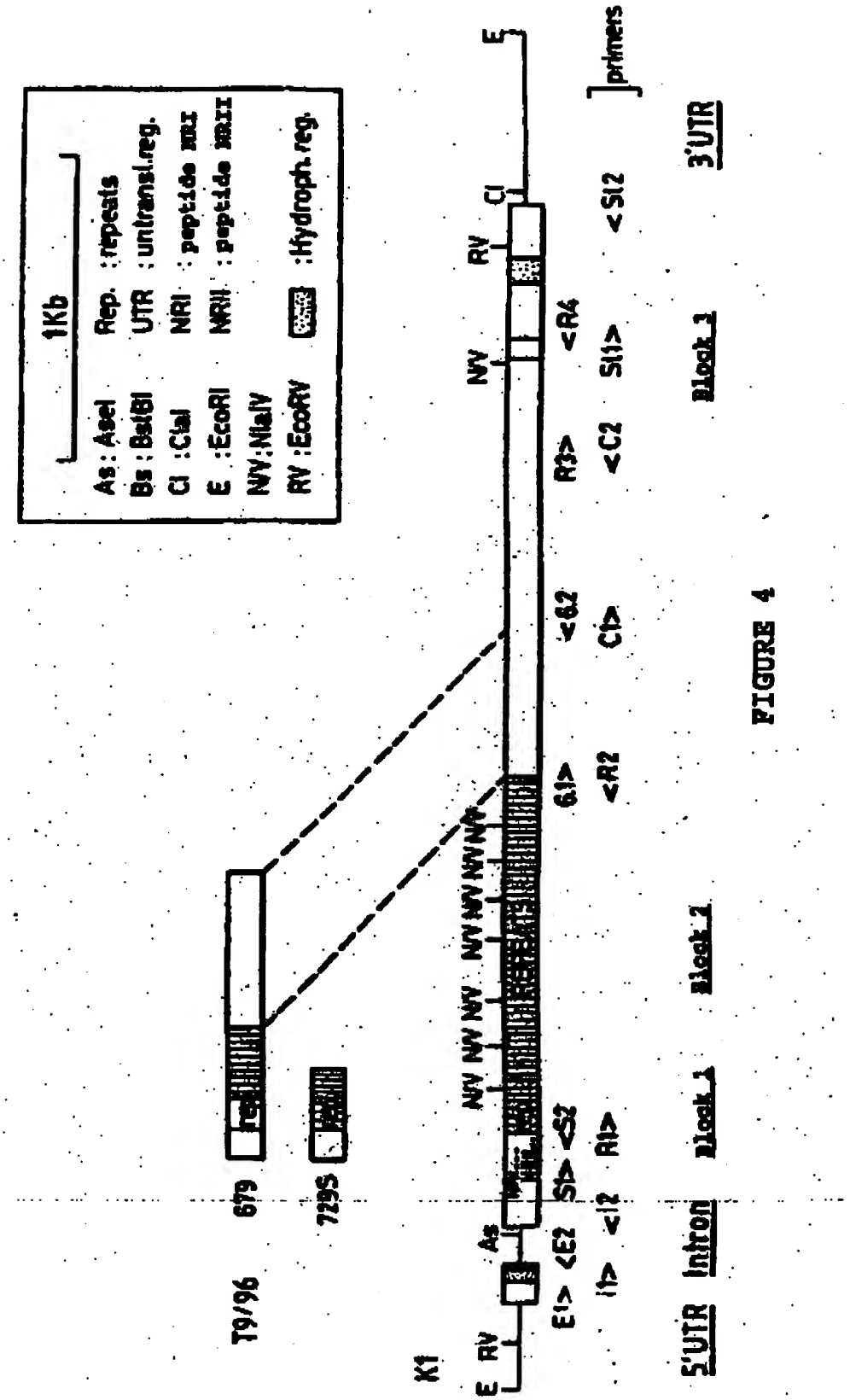

1Kb
As: AseI
Bs : BstBI
Cl :ClaI
E :EcoRI
NV:NlaIV
RV :EcoRV
Rep. :repeats
UTR : untransl.reg.
NRI : peptide NRI
NRII : peptide NRII
:Hydroph.reg.
T9/96
679
7295
K1
E RV
As
NV NV NV NV NV NV NV
E1> <E2 S1> <S2
i1> <i2 R1>
6i>
<R2
<6i2
Ci>
R3>
<C2
S11>
<S12
NV
RV
Cl
E
<R4
primers
5'UTR
Intron
Block 1
Block 2
Block 1
3'UTR

FIGURE 4

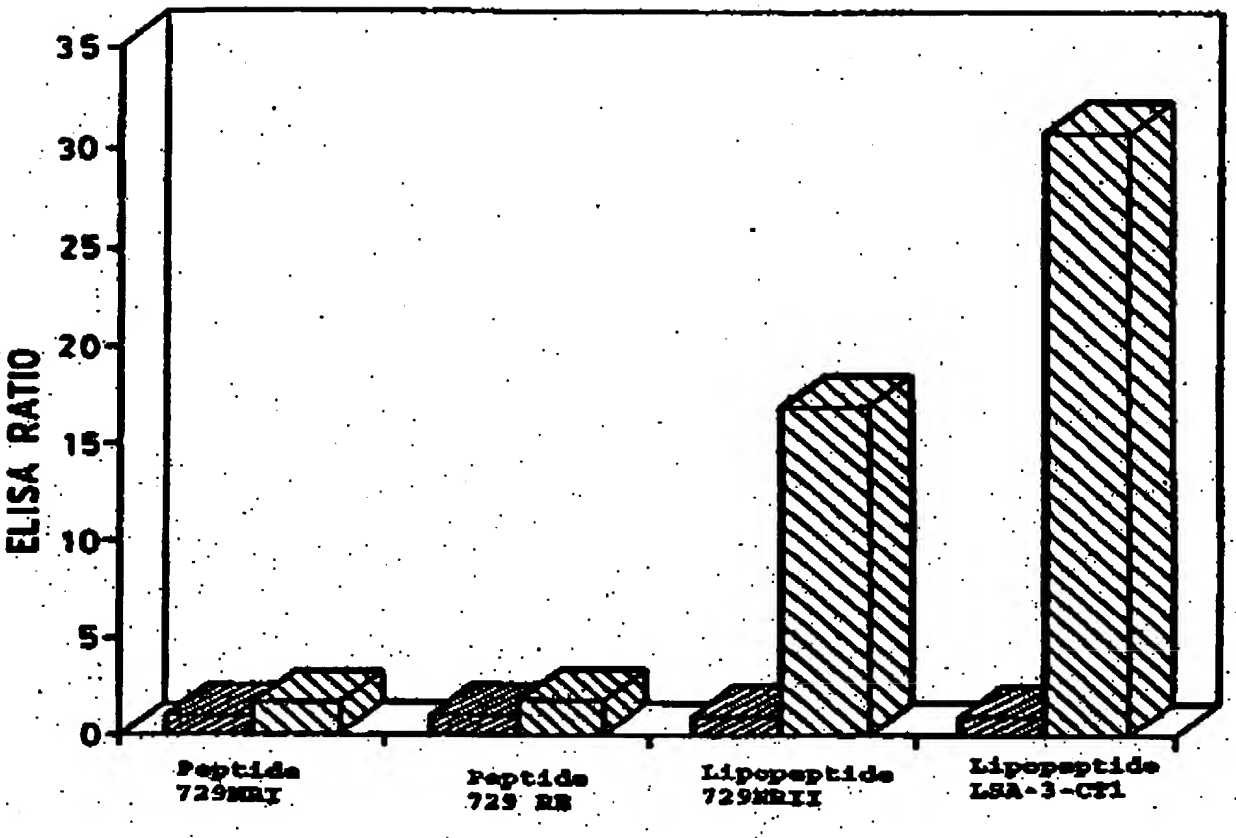

ELISA RATIO
35
30
25
20
15
10
5
0
Peptide 729MRI
Peptide 729 RR
Lipopeptide 729MRII
Lipopeptide LSA-3-CT1

FIGURE 5


Lipopeptide
Peptide
0
1
2
3
4
5
SJL/J
Balb/C
C3H/HeJ
DBA/1
C57BL/6
STRAINS

FIGURE 6

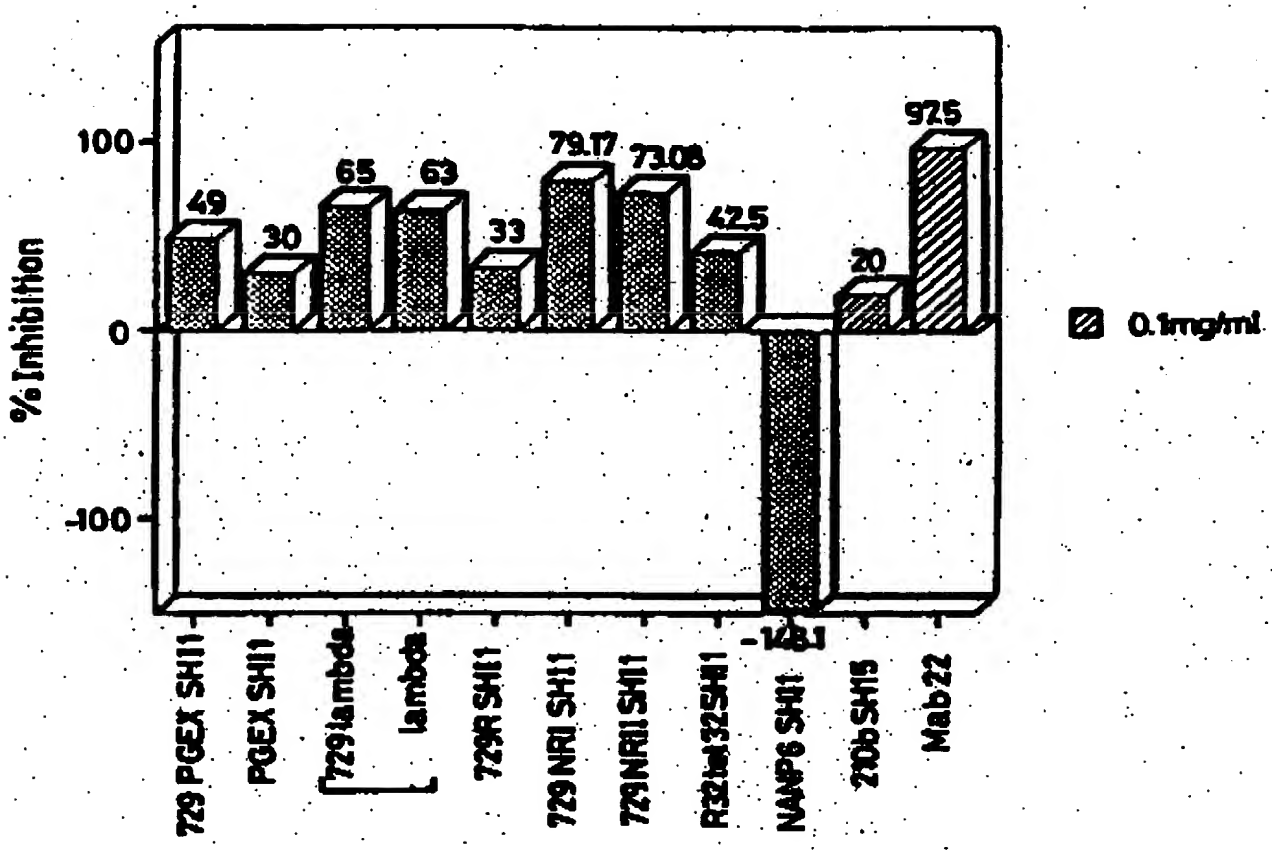

% Inhibition
100
0
-100
49
30
65
63
33
79.17
73.08
42.5
-148.1
20
97.5
729 PGEX SH11
PGEX SH11
729lambda
lambda
729R SH11
729 NR1 SH11
729NR11 SH11
R32tet32SH11
NAMP6 SH11
210b SH15
Mab22
0.1mg/ml

FIGURE 7

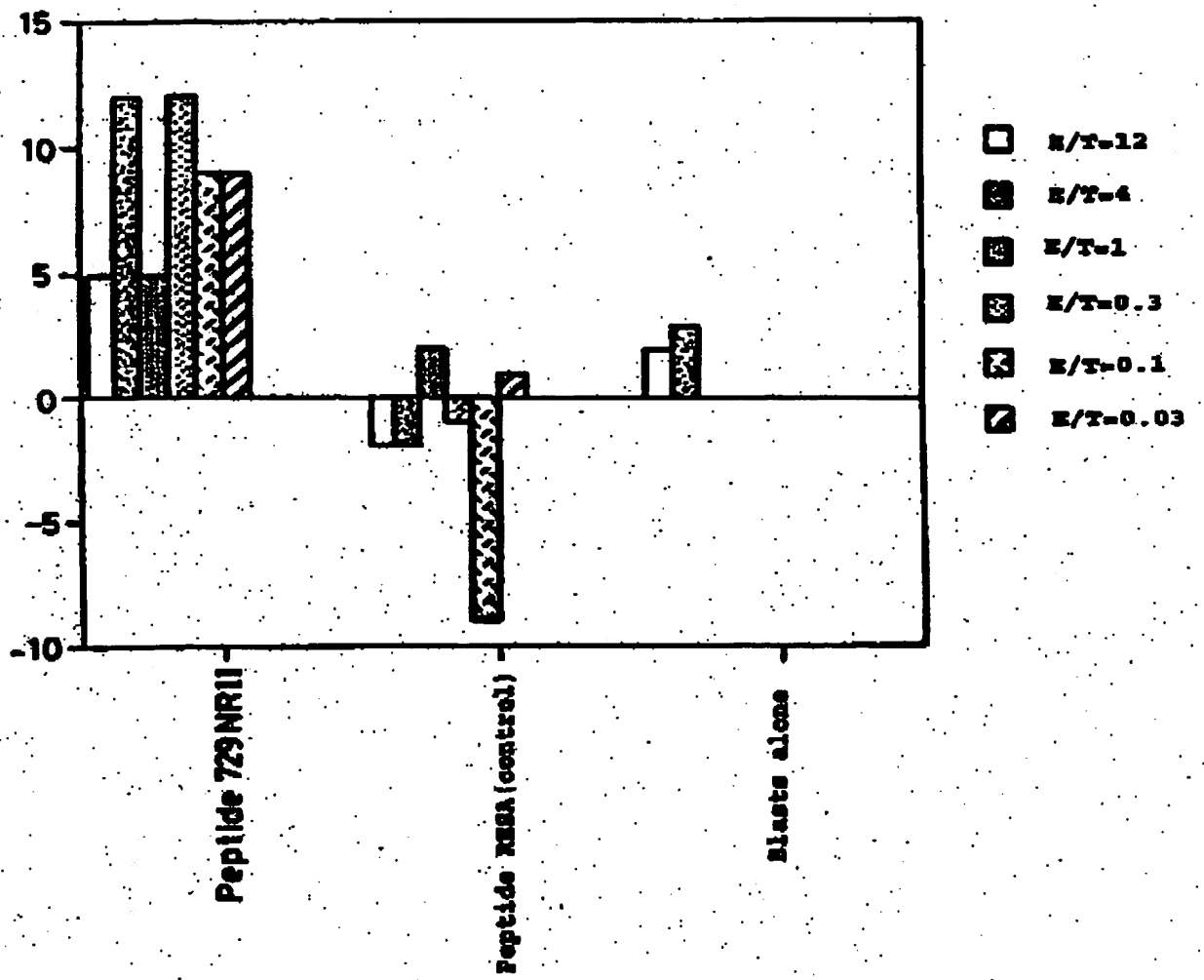

15
10
5
0
-5
-10
E/T=12
E/T=4
E/T=1
E/T=0.3
E/T=0.1
E/T=0.03
Peptide 729NR11
Peptide MBSA(control)
Blasts Alone

FIGURE 8

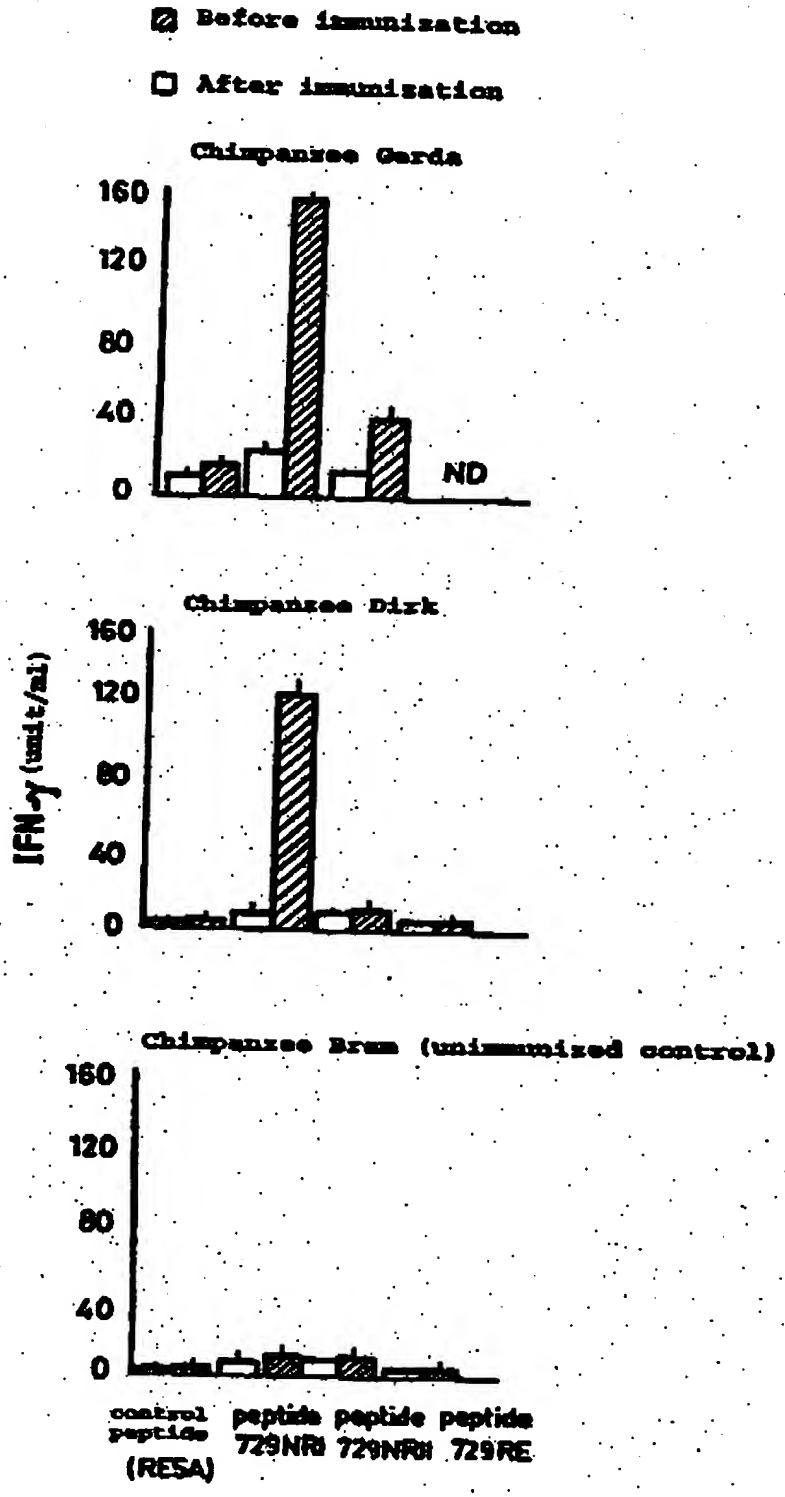

Before immunization
After immunization
Chimpanzee Garda
160
120
80
40
0
ND
Chimpanzee Dirk
IFN-γ (unit/ml)
160
120
80
40
0
Chimpanzee Brum (unimmunized control)
160
120
80
40
0
control peptide (RESA)
peptide 729NRI
peptide 729NRII
peptide 729RE

FIGURE 9